

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 16, 2005, 06:18:14 ; Search time 17058 Seconds
(without alignments)
17512.393 Million cell updates/sec

Title: US-10-791-666-1

Perfect score: 6165
Sequence: 1 atcgtgaagctcaaatatg.....gggaccagcttcagatataa 6165

Scoring table: IDENTITY NJC
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenBml:*
1: gb_ba:*
2: gb_hcg:*
3: gb_in:*
4: gb_cm:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	6165	100.0	6165	6 AR534558	Sequence
2	6165	100.0	6165	6 AX574425	Sequence
3	6161.8	99.9	6298	6 AX504254	Sequence
4	6160.2	99.9	6165	6 AX671105	Sequence
5	6160.2	99.9	8603	6 AX671108	Sequence
6	6142.2	99.6	6156	6 AX671112	Sequence
7	6126	99.4	6159	6 AX166510	Sequence
8	6121.8	99.3	6189	6 AX503780	Sequence
9	6105.4	99.0	6201	6 AX503778	Sequence
10	5843.2	94.8	8576	9 AY257469	Homo sapi
11	5837	94.7	5877	6 AR534559	Sequence
12	5837	94.7	5877	6 AR534559	Sequence
13	5661.4	91.8	6159	6 AX429514	Sequence
14	5661.4	91.8	6574	6 AX429512	Sequence
15	5105.2	82.8	6954	10 AF086824	Mus muscu
16	4909.6	79.6	5920	6 CQ723254	Sequence
17	4021.2	65.2	4967	10 AF070066	Mus muscu
18	4018.2	65.2	5019	10 MMU03904	Mus muscu
19	3906.6	63.4	5952	10 AF039218	Rattus no

20	2818.2	45.7	5261	6 AX671110	Sequence
21	2818.2	45.7	5261	9 AB023166	Homo sapi
22	2808.2	45.6	5251	6 AX925604	Sequence
23	2682.8	43.5	2896	6 CQ498283	Sequence
24	2403	39.0	3495	10 BC051165	Mus muscu
25	2403	39.0	3506	10 BC023775	Mus muscu
26	1436.2	23.3	1799	9 AY209000	Homo sapi
27	1401.4	22.7	1485	6 AX671037	Sequence
28	1401.4	22.7	1765	6 AX671044	Sequence
29	1399.8	22.7	2066	6 AX642956	Sequence
30	1398.2	22.7	1515	6 AR253937	Sequence
31	1398.2	22.7	1515	6 AR453415	Sequence
32	1373	22.3	3356	6 CQ722706	Sequence
33	1088.8	17.7	2380	6 AX671042	Sequence
34	1088.8	17.7	2380	6 AF086823	Mus muscu
35	1080.8	17.5	1386	10 AF070065	Rattus no
36	958.8	15.6	2162	6 AR542171	Sequence
37	756.4	12.3	1133	6 AR411948	Sequence
38	756.4	12.3	1133	6 AR475453	Sequence
39	754	12.2	1058	6 AX053315	Sequence
40	698	11.3	817	6 AX053416	Sequence
41	443.4	7.2	446	6 CQ483947	Sequence
42	436	7.1	485	6 CQ513776	Sequence
43	434	7.0	2284	9 AK123136	Homo sapi
44	350.4	5.7	354	6 BD059649	Secretd
45	311.2	5.0	136436	9 AC002563	Human PAC

ALIGNMENTS

RESULT 1	AR534558	6165 bp	DNA	linear	PAT 08-OCT-2004
LOCUS	AR534558				
DEFINITION	Sequence 1 from patent US 6734009.				
ACCESSION	AR534558				
VERSION	AR534558.1	GI:53924885			
KEYWORDS	Unknown.				
SOURCE	Unknown.				
ORGANISM	Unclassified.				
REFERENCE	1 (bases 1 to 6165)				
AUTHORS	Yu, X.S., Miranda, M., and Fridde, C.J.				
TITLE	Human kinases and polynucleotides encoding the same				
JOURNAL	Patent: US 6734009-A 1 11-MAY-2004;				
FEATURES	location/Qualifiers				
source	1..6165				
	/organism="unknown"				
	/mol_type="genomic DNA"				

ORIGIN

Query Match	100.0%;	Score 6165;	DB 6;	Length 6165;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 6165;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	ATGTTGAAGTTCAATATGAGCGCGGAATCCTTTGATGCTGCTGAACCAT	60	
DB	1	ATGTTGAAGTTCAATATGAGCGCGGAATCCTTTGATGCTGCTGAACCAT	60	
QY	61	GCGAGCGGCGCTTCAGGCTGAATCTGTTCTTCAGGGGAAACACCTTTATGACTCA	120	
DB	61	GCGAGCGGCGCTTCAGGCTGAATCTGTTCTTCAGGGGAAACACCTTTATGACTCA	120	
QY	121	CAGCAGATGTCCTCTCTTCCCGAAGGAGATTTAATGCCCTTTGTTCTTTGAA	180	
DB	121	CAGCAGATGTCCTCTCTTCCCGAAGGAGATTTAATGCCCTTTGTTCTTTGAA	180	
QY	181	GAATGACGAGCTGCTGCTGATGAAGATTAAAGCACTGACCACTTTGTCGGAGATAT	240	
DB	181	GAATGACGAGCTGCTGCTGATGAAGATTAAAGCACTGACCACTTTGTCGGAGATAT	240	
QY	241	TCCGACCATTAAGTGAATTAAGAGAGCTTCAGGCTTCGCAAGGACCTTGAAAGTCA	300	

Db 241 TCCGACACATAGCTGAGTTACAGAGCTCCAGCCTTCGCAAGACCTTCGAACTGACA 300
Qy 301 AGCTTTAGAGTTTGCTGCTCACTTTGCTGAAGTGCAGGTGGTAAAGAGAAACCAACGGG 360
Db 301 AGCTTTAGAGTTTGCTGCTCACTTTGCTGAAGTGCAGGTGGTAAAGAGAAACCAACGGG 360
Qy 361 GACATCTATCTATGAAGTGAAGAGAAAGGCTTTATTTGGCCCGAGAGCAGTTTCA 420
Db 361 GACATCTATCTATGAAGTGAAGAGAAAGGCTTTATTTGGCCCGAGAGCAGTTTCA 420
Qy 421 TTTTGTGAGGAAGCGGAAATATATCTCGAAGCACAAGCCCGTGAATCCCCAAATTA 480
Db 421 TTTTGTGAGGAAGCGGAAATATATCTCGAAGCACAAGCCCGTGAATCCCCAAATTA 480
Qy 481 CAGTATGCCCTTTGAGGACAAAATACCTTTATTTGGTCAATGAATTAACGCTGAGGG 540
Db 481 CAGTATGCCCTTTGAGGACAAAATACCTTTATTTGGTCAATGAATTAACGCTGAGGG 540
Qy 541 GACTTGCTGCTCACTTTGAAATGATGAGGACAGTTAGATGAAAACCTGATACATTT 600
Db 541 GACTTGCTGCTCACTTTGAAATGATGAGGACAGTTAGATGAAAACCTGATACATTT 600
Qy 601 TACCTAGCTGAGTGAATTTTGCTGTTCAAGGCTTCATCTGAATGAGATACGTCATCGA 660
Db 601 TACCTAGCTGAGTGAATTTTGCTGTTCAAGGCTTCATCTGAATGAGATACGTCATCGA 660
Qy 661 GAATCATAGCTGAGAAATCTCTGTTGACCGCAAGAGCAATCAAGCTGAGATTTT 720
Db 661 GAATCATAGCTGAGAAATCTCTGTTGACCGCAAGAGCAATCAAGCTGAGATTTT 720
Qy 721 GGAATCTGCGCGAAATGAATTAACAAGATGGTGAATGCCAACTCCGATTTGGGACC 780
Db 721 GGAATCTGCGCGAAATGAATTAACAAGATGGTGAATGCCAACTCCGATTTGGGACC 780
Qy 781 CCAGATTACATGAGCTCTCTGAAGTCTGACTGTGATGAACGGGATGGAAGGACACTAC 840
Db 781 CCAGATTACATGAGCTCTCTGAAGTCTGACTGTGATGAACGGGATGGAAGGACACTAC 840
Qy 841 GGCCTGAGCTGTGACTGTGAGTGTGAGTGGGCTGATGGCTTATGATGATTTATGGGAGA 900
Db 841 GGCCTGAGCTGTGACTGTGAGTGTGAGTGGGCTGATGGCTTATGATGATTTATGGGAGA 900
Qy 901 TCCCTCTGCGAGAGGAACTCTGCAGAACCTTAATTAACATTAATTTCCAGCGG 960
Db 901 TCCCTCTGCGAGAGGAACTCTGCAGAACCTTAATTAACATTAATTTCCAGCGG 960
Qy 961 TTTTGAATTTTCAGATGACCCCAAGTGAAGTGAATCTTTGATGATTCGATTCGAAAG 1020
Db 961 TTTTGAATTTTCAGATGACCCCAAGTGAAGTGAATCTTTGATGATTCGATTCGAAAG 1020
Qy 1021 TTGTTGTGCGGCGAAGAGAGACTGAAGTGAAGGTTGATGATGATTCGATTCGAAAG 1080
Db 1021 TTGTTGTGCGGCGAAGAGAGACTGAAGTGAAGGTTGATGATGATTCGATTCGAAAG 1080
Qy 1081 TCTAAATTTGATGAGAAACATTTGTAATCTCTCCCTCCCTCCCTCCCTCCCTCCAG 1140
Db 1081 TCTAAATTTGATGAGAAACATTTGTAATCTCTCCCTCCCTCCCTCCCTCCCTCCAG 1140
Qy 1141 TCTGACATGACCTTCATTTTGTGATGACAGAGAAATTTGAGGTTTTCATCTCT 1200
Db 1141 TCTGACATGACCTTCATTTTGTGATGACAGAGAAATTTGAGGTTTTCATCTCT 1200
Qy 1201 CCGTGCAGCTGAGCCCTCAGGCTTCTCGGGTGAAGAACTGCCGTTTGGGGTTTTCG 1260
Db 1201 CCGTGCAGCTGAGCCCTCAGGCTTCTCGGGTGAAGAACTGCCGTTTGGGGTTTTCG 1260
Qy 1261 TACAGCAAGGCACTGGGGAATTTTGTGATGATCTGATCTGTTGTGCTGGGTTTCG 1320
Db 1261 TACAGCAAGGCACTGGGGAATTTTGTGATGATCTGATCTGTTGTGCTGGGTTTCG 1320
Qy 1321 CCTGCAGAGCTAGCTCCATGAGAAAGAACTTTCATCAAAAGCAAGAGCTACAAAGC 1380
Db 1321 CCTGCAGAGCTAGCTCCATGAGAAAGAACTTTCATCAAAAGCAAGAGCTACAAAGC 1380

Qy 1381 TCTCAGGACAAAGTGTCAAAAGATGAGACAGGAAATGACCCGGTTACATCGGAGTGTCA 1440
Db 1381 TCTCAGGACAAAGTGTCAAAAGATGAGACAGGAAATGACCCGGTTACATCGGAGTGTCA 1440
Qy 1441 GAGGTGAGGCTGTGCTTACTGAGAAAGAGTGAAGTGAAGGCTCTGAGACTCAGAGA 1500
Db 1441 GAGGTGAGGCTGTGCTTACTGAGAAAGAGTGAAGTGAAGGCTCTGAGACTCAGAGA 1500
Qy 1501 TCCCTCTGAGAGAGGACCTTGTCTACCTACATCAAGAAATGACAGTACCTTAAAGCGAAGT 1560
Db 1501 TCCCTCTGAGAGAGGACCTTGTCTACCTACATCAAGAAATGACAGTACCTTAAAGCGAAGT 1560
Qy 1561 TTGAGCAAGCACGATGAGGTTGTCCAGAGAGATGACAAAGCACTGACCTTCTCAT 1620
Db 1561 TTGAGCAAGCACGATGAGGTTGTCCAGAGAGATGACAAAGCACTGACCTTCTCAT 1620
Qy 1621 GATATCAGAGAGAGAGCCCGGAAAGCTCCAGAAATCAAAAGACAGAGTCCAGGCTCAA 1680
Db 1621 GATATCAGAGAGAGAGCCCGGAAAGCTCCAGAAATCAAAAGACAGAGTCCAGGCTCAA 1680
Qy 1681 GTGGAAGAAATGAGTGAATGATGAATCAGTTGGAAGAGGATCTTGTCTCAGCAAGAA 1740
Db 1681 GTGGAAGAAATGAGTGAATGATGAATCAGTTGGAAGAGGATCTTGTCTCAGCAAGAA 1740
Qy 1741 CCGAGTATCTTACGAATCTGAGCTGAGAGAGTCTCGGCTTCTGCTGAAGAAATTCAG 1800
Db 1741 CCGAGTATCTTACGAATCTGAGCTGAGAGAGTCTCGGCTTCTGCTGAAGAAATTCAG 1800
Qy 1801 CCGAAGCGACAGAAATGTCAGCAATTAATGTAAGCTTAAGATCAAGGAGACCTGAA 1860
Db 1801 CCGAAGCGACAGAAATGTCAGCAATTAATGTAAGCTTAAGATCAAGGAGACCTGAA 1860
Qy 1861 GTGGGAAATATGCGAAATGCGAAGATCAATGCTGACAGAGCTCAAAATTCAGAGAG 1920
Db 1861 GTGGGAAATATGCGAAGATGCGAAGATCAATGCTGACAGAGCTCAAAATTCAGAGAG 1920
Qy 1921 CTCCAAGAGAACTGAGAGAGGCTGTAAAGCCAGACCGAAGGCTGCTGACAG 1980
Db 1921 CTCCAAGAGAACTGAGAGAGGCTGTAAAGCCAGACCGAAGGCTGCTGACAG 1980
Qy 1981 AATATCCGCGAGCAAGAGAGGAGCGGAGAGAGCTGAGAAAGCTGCAGAACCGAGAG 2040
Db 1981 AATATCCGCGAGCAAGAGAGGAGCGGAGAGAGCTGAGAAAGCTGCAGAACCGAGAG 2040
Qy 2041 GATTCTTCTGAAGGCTCAGAAAGAGCTGTGGAAGCTGAGGAAGCGCCCATTTCTTG 2100
Db 2041 GATTCTTCTGAAGGCTCAGAAAGAGCTGTGGAAGCTGAGGAAGCGCCCATTTCTTG 2100
Qy 2101 GAGAACAAAGTAAAGAGACTAGAGACATGAGCGTGAAGAAACAGACTGAAGATGAC 2160
Db 2101 GAGAACAAAGTAAAGAGACTAGAGACATGAGCGTGAAGAAACAGACTGAAGATGAC 2160
Qy 2161 ATCCAGACAAATCCCAACAGATCCAGAGATGCTGATTAATAATTTGAGGCTCCAGAG 2220
Db 2161 ATCCAGACAAATCCCAACAGATCCAGAGATGCTGATTAATAATTTGAGGCTCCAGAG 2220
Qy 2221 AAACATCGGAGGCGCAAGTCTCAGCCGAGCACTAGAAAGTGCACCTGAAACAGAAAGAG 2280
Db 2221 AAACATCGGAGGCGCAAGTCTCAGCCGAGCACTAGAAAGTGCACCTGAAACAGAAAGAG 2280
Qy 2281 CAGCACTATGAGAGAAAGATTAAGTGTGGAATCAATCAGATTAAGAAAGACCTGCTGAC 2340
Db 2281 CAGCACTATGAGAGAAAGATTAAGTGTGGAATCAATCAGATTAAGAAAGACCTGCTGAC 2340
Qy 2341 AAGAGACACTGAGAGAACTGATGAGACACAGAGAGAGGCGCCATGAGAGAGGCAAA 2400
Db 2341 AAGAGACACTGAGAGAACTGATGAGACACAGAGAGAGGCGCCATGAGAGAGGCAAA 2400
Qy 2401 ATTCTCAGCGAACAGAGCGATGATCAATGATTAAGATTCGAAGATCAAGATCCCTGGA 2460
Db 2401 ATTCTCAGCGAACAGAGCGATGATCAATGATTAAGATTCGAAGATCAAGATCCCTGGA 2460

Qy	2461	CAGAGGATTGGGAACCTGTCTGAAAGCCATAATTAACCTTGACGAATATAGCAAGCTCTTTTACC	2520
Dp	2461	CAGAGGATTGGGAACCTGTCTGAAAGCCATAATTAACCTTGACGAATATAGCAAGCTCTTTTACC	2520
Qy	2521	CAAAAGGAAACAAGAAAGGCCCAAGAAAGAGATGATTTCTGAACTCAGCAACAGAAATTTTAC	2580
Dp	2521	CAAAAGGAAACAAGAAAGGCCCAAGAAAGATGATTTCTGAACTCAGCAACAGAAATTTTAC	2580
Qy	2581	CTGGAGACAACAGGCTGGGAAGTTTGGAGGCCAGAACCCGAAACCTGGAGAGCAGCTGGAG	2640
Dp	2581	CTGGAGACAACAGGCTGGGAAGTTTGGAGGCCAGAACCCGAAACCTGGAGAGAGCAGCTGGAG	2640
Qy	2641	AAGATCAAGCCACAAGACCAACAGTGAACAAGATCGGCTGCTGGAACCTGGAGACAAGATTG	2700
Dp	2641	AAGATCAAGCCACAAGACCAACAGTGAACAAGATCGGCTGCTGGAACCTGGAGACAAGATTG	2700
Qy	2701	CGGAGAGTCACTTAAGAGCAAGAGAGAGAACTGAGAGCTCAAGCCCACTCAACAG	2760
Dp	2701	CGGAGAGTCACTTAAGAGCAAGAGAGAGAAACTGAGAGCTCAAGCCCACTCAACAG	2760
Qy	2761	CTACAGCTCTCCCTGCAAGAGCCGAGTCAACAGTTGACAGCCCTGCAAGGCTGCACGGGCG	2820
Dp	2761	CTACAGCTCTCCCTGCAAGAGCCGAGTCAACAGTTGACAGCCCTGCAAGGCTGCACGGGCG	2820
Qy	2821	GCCTTGGAGAGCCAGCTTCGCGAGGCGAAGACAGAGCTGGAGAGACCAACAGAAAGCT	2880
Dp	2821	GCCTTGGAGAGCCAGCTTCGCGAGGCGAAGACAGAGCTGGAGAGACCAACAGAGCT	2880
Qy	2881	GAAAGAGAGATCCAGGCACTCAACGGCACTGAGATGAATCCAGCCCAATTTGATGCT	2940
Dp	2881	GAAAGAGAGATCCAGGCACTCAACGGCACTGAGATGAATCCAGCCCAATTTGATGCT	2940
Qy	2941	CTTCGTAAACAGCTGTACTGTAAATACAGACCTTGGAGAGCAGCTAAACAGCTGACCGAG	3000
Dp	2941	CTTCGTAAACAGCTGTACTGTAAATACAGACCTTGGAGAGCAGCTAAACAGCTGACCGAG	3000
Qy	3001	GACAAACGCTGAACCTCAACCAAACTTCTAATTGTCAAACAACTCGATGAGGCTTCT	3060
Dp	3001	GACAAACGCTGAACCTCAACCAAACTTCTAATTGTCAAACAACTCGATGAGGCTTCT	3060
Qy	3061	GGCCCCAAACGACGAGATTGTAACACTCGGAAGTAAGTGAACATTCGCCCGGGAATC	3120
Dp	3061	GGCCCCAAACGACGAGATTGTAACACTCGGAAGTAAGTGAACATTCGCCCGGGAATC	3120
Qy	3121	ACGGAACGAGAGATGACAGCTTACACAGCCAGAAACGATGAGGCTCTGAAACACAGC	3180
Dp	3121	ACGGAACGAGAGATGACAGCTTACACAGCCAGAAACGATGAGGCTCTGAAACACAGC	3180
Qy	3181	TGCACCATGCTGAGAGAAACAGTCAATGAGATTGGAGGCCCTTAAACGATGAGCTCTAGAA	3240
Dp	3181	TGCACCATGCTGAGAGAAACAGTCAATGAGATTGGAGGCCCTTAAACGATGAGCTCTAGAA	3240
Qy	3241	AAAAGCGGCACTGGGAGGCTCTGGAGAGAGCTCTGGGTGATGAGAAATCCAGTTTGAG	3300
Dp	3241	AAAAGCGGCACTGGGAGGCTCTGGAGAGAGCTCTGGGTGATGAGAAATCCAGTTTGAG	3300
Qy	3301	TGTGGGGTTCAGAGAGCTGCAAGAGATGCTGACACCGAGAAACAGAGCAGGCGAGAGCC	3360
Dp	3301	TGTGGGGTTCAGAGAGCTGCAAGAGATGCTGACACCGAGAAACAGAGCAGGCGAGAGCC	3360
Qy	3361	GATCAGCGGATCAACCGAGTCTCGCCAGGTGGTGAAGCTGCGAGTGAAGAGCAACAGCT	3420
Dp	3361	GATCAGCGGATCAACCGAGTCTCGCCAGGTGGTGAAGCTGCGAGTGAAGAGCAACAGCT	3420
Qy	3421	GAGATTCTCGCTCTGCAAGAGCTCTCAAAAGCAGAAAGCTGAAGGCCAGAGGCTCTCT	3480
Dp	3421	GAGATTCTCGCTCTGCAAGAGCTCTCAAAAGCAGAAAGCTGAAGGCCAGAGGCTCTCT	3480
Qy	3481	GACAAGCTCAATGACCTGGAGAGAAAGCATGCTATGCTGAAATGAATGCCCAAGCTTA	3540
Dp	3481	GACAAGCTCAATGACCTGGAGAGAAAGCATGCTATGCTGAAATGAATGCCCAAGCTTA	3540
Qy	3541	CAGACGAAGCTGAGAGCTGAACGAAGCTCAACAGAGGCTTCTGGAAGAGCAAGCCAAA	3600

Db	3541	CAGCAGAAAGCTGAGACTGAAAGAGCTCAACAGAGCTTCTGGAAAGGCAAGCCAAA	3600
Qy	3601	TTACAGCAGCAGATGGAACCTCGAGAAAATTCATTTCCGTCCTGACTCAGAGACTGCAA	3660
Db	3601	TTACAGCAGCAGATGGAACCTCGAGAAAATTCATTTCCGTCCTGACTCAGAGACTGCAA	3660
Qy	3661	GAAAGCTTAGATCGGAGCTGATCTACTGAAGACAGAAAGAAAGTGAACCTTGAGCTG	3720
Db	3661	GAAAGCTTAGATCGGAGCTGATCTACTGAAGACAGAAAGAAAGTGAACCTTGAGCTG	3720
Qy	3721	GAAAAATTCAAGTTCCTATTCATGAAAAGGTAAGAAAATGAAAGCACTATTTCTCAA	3780
Db	3781	CAAAACCAACCTATTGATTTTCTGCAAGCAAAATGACCAACCTGCTAAAAAGAAAAG	3840
Qy	3781	CAAAACCAACCTATTGATTTTCTGCAAGCAAAATGACCAACCTGCTAAAAAGAAAAG	3840
Db	3841	GTTCCCTGCACTAGTACATGAGCTGAAGCTGGCCCTGGAAGAGAAAGCTCGCTGTGCA	3900
Qy	3841	GTTCCCTGCACTAGTACATGAGCTGAAGCTGGCCCTGGAAGAGAAAGCTCGCTGTGCA	3900
Db	3901	GAGCTAGAGAGAGCCCTTCAGAAAGCCGCAATCGAGCTCCGCTCCGCGGAGAGAACT	3960
Qy	3901	GAGCTAGAGAGAGCCCTTCAGAAAGCCGCAATCGAGCTCCGCTCCGCGGAGAGAACT	3960
Db	3961	GCCCAACCGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG	4020
Qy	3961	GCCCAACCGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG	4020
Db	4021	ATGCGCATGTCGCGCATCGTGGCGGTGCGCAAGACACAGCCAGCCAGTGCATGAGCTGTG	4080
Qy	4021	ATGCGCATGTCGCGCATCGTGGCGGTGCGCAAGACACAGCCAGCCAGTGCATGAGCTGTG	4080
Db	4081	GCCCCCGCCATCCAGCCCGCAGAAAGAGTCTTCACTCCAGAGAAATTTAGTCGGCGCTT	4140
Qy	4081	GCCCCCGCCATCCAGCCCGCAGAAAGAGTCTTCACTCCAGAGAAATTTAGTCGGCGCTT	4140
Db	4141	AAGGAACGCATGACACACAAATTTCTCAACCGATTCAAGTGAAGACTGAAATGCGAGCC	4200
Qy	4141	AAGGAACGCATGACACACAAATTTCTCAACCGATTCAAGTGAAGACTGAAATGCGAGCC	4200
Db	4201	ACAAAGTGTGCTGT	4260
Qy	4201	ACAAAGTGTGCTGT	4260
Db	4261	GAAATGCAAGTATGT	4320
Qy	4261	GAAATGCAAGTATGT	4320
Db	4321	CCTGCGAATATGCAACACTTCCAGAGGCTTGTGCGGTGCAAAATGAACTCCCA	4380
Qy	4321	CCTGCGAATATGCAACACTTCCAGAGGCTTGTGCGGTGCAAAATGAACTCCCA	4380
Db	4381	GCTTCAGAACCAAGAGCCAGCAGACCTTGTGCACTGGAAGGTGTGATGAAAGTGTCCC	4440
Qy	4381	GCTTCAGAACCAAGAGCCAGCAGACCTTGTGCACTGGAAGGTGTGATGAAAGTGTCCC	4440
Db	4441	AGGAATTAACAAACGAGAGCAGCAAGCTGTGGACAGGAAGTCACTTGTCTGAGGGATCA	4500
Qy	4441	AGGAATTAACAAACGAGAGCAGCAAGCTGTGGACAGGAAGTCACTTGTCTGAGGGATCA	4500
Db	4501	AAAGTCCCTATTATGACAAATGAAGCAGAAAGCTGTGACAGAGGCCGATGGAAGATTT	4560
Qy	4501	AAAGTCCCTATTATGACAAATGAAGCAGAAAGCTGTGACAGAGGCCGATGGAAGATTT	4560
Db	4561	GAGCTGTGCTTCCGACCGGAGATGTATCTATTCATGTGTGCGTGTGTGTGTGTGTGT	4620
Qy	4561	GAGCTGTGCTTCCGACCGGAGATGTATCTATTCATGTGTGCGTGTGTGTGTGTGTGT	4620
Db	4621	GCAATTAACGCAAGCAGATGTCCCATTCATCTGAAGATGGAATCTCACCCGCAACC	4680
Qy	4621	GCAATTAACGCAAGCAGATGTCCCATTCATCTGAAGATGGAATCTCACCCGCAACC	4680

Db	4621	GGAAATACAGCCAAAGACGATGTCCCATACATACGAAAGATGAATCTCACCCCGCACCC	4680
Qy	4681	ACCTGCTGCCCCGGAGAAACCTCTACTTGTGTAAGCTCCAGCTTCCCTGACAAACAGCGC	4740
Db	4681	ACCTGCTGCCCCGGAGAAACCTCTACTTGTGTAAGCTCCAGCTTCCCTGACAAACAGCGC	4740
Qy	4741	TGGGTCAACCCGCTTGAAATCAAGTGTCCGAGGTGGAGAGATTCTTAGGGAAAAAGCAGAA	4800
Db	4741	TGGGTCAACCCGCTTGAAATCAAGTGTCCGAGGTGGAGAGATTCTTAGGGAAAAAGCAGAA	4800
Qy	4801	GCTGATGCTAAACCTGCTGAAACCTCCGTCGAAACCTGGCGAAACCTGAAAGGTATGACCCGCTAGAC	4860
Db	4801	GCTGATGCTAAACCTGCTGAAACCTCCGTCGAAACCTGGCGAAACCTGAAAGGTATGACCCGCTAGAC	4860
Qy	4861	ATGAACCTGACGCTGCCCCCTTCAAGTACCAAGGTGATGTGTGGGACCCGAGAGAGGCTC	4920
Db	4861	ATGAACCTGACGCTGCCCCCTTCAAGTACCAAGGTGATGTGTGGGACCCGAGAGAGGCTC	4920
Qy	4921	TACGCCCCGTAATGTCTTGAAAAACCTCCCTAAACCAATGCTCCAGAAATTGGAGCAGTCTTC	4980
Db	4921	TACGCCCCGTAATGTCTTGAAAAACCTCCCTAAACCAATGCTCCAGAAATTGGAGCAGTCTTC	4980
Qy	4981	CAAAATTATATATCAAGAGACCTGGAGAAAGTACTCATCATATAGAGAGAGAAAGACCGGGCA	5040
Db	4981	CAAAATTATATATCAAGAGACCTGGAGAAAGTACTCATCATATAGAGAGAGAAAGACCGGGCA	5040
Qy	5041	CTGTGTCTTGTGACGCTGAAAGAAAGTGAACAGTCCCTGGCCCAAGTCCCACTTGCTGGCC	5100
Db	5041	CTGTGTCTTGTGACGCTGAAAGAAAGTGAACAGTCCCTGGCCCAAGTCCCACTTGCTGGCC	5100
Qy	5101	CAGCCCGACATCTCAACCCAAATTTTGAAGCTGTCAAGGGCTGCCACTTGTTTGGGGCA	5160
Db	5101	CAGCCCGACATCTCAACCCAAATTTTGAAGCTGTCAAGGGCTGCCACTTGTTTGGGGCA	5160
Qy	5161	GGCAAGATGTGAAGCGGGCTGTGATCTGTGACGACCAATGCCACGAAAGTCATCTTC	5220
Db	5161	GGCAAGATGTGAAGCGGGCTGTGATCTGTGACGACCAATGCCACGAAAGTCATCTTC	5220
Qy	5221	CGCTACACGAAAACTCAGCAAAATACGTGATCCGGAAGAGATAGACCTCAGAGCCC	5280
Db	5221	CGCTACACGAAAACTCAGCAAAATACGTGATCCGGAAGAGATAGACCTCAGAGCCC	5280
Qy	5281	TGCAGCTGTATCTCACTTCACCAATTAACGATATCTCTCATTTGGAAACAAATAATCTACGAA	5340
Db	5281	TGCAGCTGTATCTCACTTCACCAATTAACGATATCTCTCATTTGGAAACAAATAATCTACGAA	5340
Qy	5341	ATGCACATGAAGCAGTACAGCTGCGAAGAAATTCCTGGATAGAAATGACCAATTCCTGGCA	5400
Db	5341	ATGCACATGAAGCAGTACAGCTGCGAAGAAATTCCTGGATAGAAATGACCAATTCCTGGCA	5400
Qy	5401	CTTGCTGTGTGTGGCGGCTCTTCCACAGAGCTCCCTGTCCAATTCGACAGGTGAACAGC	5460
Db	5401	CTTGCTGTGTGTGGCGGCTCTTCCACAGAGCTCCCTGTCCAATTCGACAGGTGAACAGC	5460
Qy	5461	GCAAGGCAAGCAAGAGAGTACTTGCTGTGTTCACGAATTTGAGAGTTCGTGGAATTC	5520
Db	5461	GCAAGGCAAGCAAGAGAGTACTTGCTGTGTTCACGAATTTGAGAGTTCGTGGAATTC	5520
Qy	5521	TAGGGAAGAGTGAAGCGGCAACAGACATCTCAAGTGAAGTGCCTTACCTTGGGCTTGGC	5580
Db	5521	TAGGGAAGAGTGAAGCGGCAACAGACATCTCAAGTGAAGTGCCTTACCTTGGGCTTGGC	5580
Qy	5581	TACAGAGAACCTTATCTGTGTGTGACCACTTCAACATCATCTGAAAGTAAATTGATCCAG	5640
Db	5581	TACAGAGAACCTTATCTGTGTGTGACCACTTCAACATCATCTGAAAGTAAATTGATCCAG	5640
Qy	5641	GCAAGCTCTTACAGAGGAGACCCCTGCCCCGAGCGTACCTGACATCCCGAACCCGCGCTAC	5700
Db	5641	GCAAGCTCTTACAGAGGAGACCCCTGCCCCGAGCGTACCTGACATCCCGAACCCGCGCTAC	5700
Qy	5701	CTGGGCCCCCTCCATTTCTCTAGAGAGGATTACTTGGCGCTCTCATACAGATTAATTA	5760
Db	5701	CTGGGCCCCCTCCATTTCTCTAGAGAGGATTACTTGGCGCTCTCATACAGATTAATTA	5760

Qy	5761	AGGGGTCATTTGCTGCAAGGAAACCTCGTAAAGAGTCCGGCACTGAAACACCAACCGGGGC	5820
Db	5761	AGGGTCATTTGCTGCAAGGAAACCTCGTAAAGAGTCCGGCACTGAAACACCAACCGGGGC	5820
Qy	5821	CCGTGCACTTCCCGCAGAGCCCAACAAGGAGAGCCACCCAGCTCAACAGCAATC	5880
Db	5821	CCGTGCACTTCCCGCAGAGCCCAACAAGGAGAGCCCACTCACTCAACAGCAATC	5880
Qy	5881	ACCAAGCGGTGGCTTCAGGCCCAAGCGCCGCGCAAGGCCCAACCGCGAGAGCA	5940
Db	5881	ACCAAGCGGTGGCTTCAGGCCCAAGCGCCGCGCAAGGCCCAACCGCGAGAGCA	5940
Qy	5941	AGCAACCCCAACCGCTACCGCGAAGGGCGGACCGAGCTGCGCAGAGGACAACTCTCTGGC	6000
Db	5941	AGCAACCCCAACCGCTACCGCGAAGGGCGGACCGAGCTGCGCAGAGGACAACTCTCTGGC	6000
Qy	6001	CGCCCCCTGAGCGAGAGAAATCCCGCGCCGAGATCTCAGCAGCGGAGAGAGCGGTCC	6060
Db	6001	CGCCCCCTGAGCGAGAGAAATCCCGCGCCGAGATCTCAGCAGCGGAGAGAGCGGTCC	6060
Qy	6061	CCCGGAGGCTGTTGAAAGACAGCAGCAGGAGCGCGCTGCTGCGGAGCGCTGAGAGCC	6120
Db	6061	CCCGGAGGCTGTTGAAAGACAGCAGCAGGAGCGCGCTGCTGCGGAGCGCTGAGAGCC	6120
Qy	6121	CCGCTGTCTCCAGGTGAACAAAGTCTTGAGACCAAGTCTTCAATATTA	6165
Db	6121	CCGCTGTCTCCAGGTGAACAAAGTCTTGAGACCAAGTCTTCAATATTA	6165

RESULT 2				
AXS74425				
LOCUS	AXS74425	6165 bp	DNA	linear
DEFINITION	Sequence	1 from Patent WO02059325.		
ACCESSION	AXS74425			
VERSION	AXS74425.1	GI:27551751		
KEYWORDS				
SOURCE				
ORGANISM	Homo sapiens (human)			
	Homo sapiens			
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
REFERENCE	1			
AUTHORS	Yu,X., Miranda,M. and Friddle,C.J.			
TITLE	Human kinases and polynucleotides encoding the same			
JOURNAL	Patent: WO 02059325-A 1 01-AUG-2002;			
FEATURES	Lexicon Genetics Incorporated (US)			
	Location/Qualifiers			
source	1..6165			
	/organism="Homo sapiens"			
	/mol_type="unassigned DNA"			
	/db_xref="taxon:9606"			
ORIGIN				

Query Match	Best Local Similarity	100.0%;	Score 6165;	DB 6;	Length 6165;
Match 6165;	Conservative	100.0%;	Pred. No. 0;	Mismatches 0;	Indels 0;
					Gaps 0;
Qy	1	ATGTTGAAGTTCAAAATATGAGCGCGGAATCCTTTGATGCTGTCTGCTGAACCATTT	60		
Db	1	ATGTTGAAGTTCAAAATATGAGCGCGGAATCCTTTGATGCTGTCTGCTGAACCATTT	60		
Qy	61	GCCACGCCGGGCGCTTCAGGCTGAATCTGTTCTTCCAGGGGAAACCACTTTATGACTCAA	120		
Db	61	GCCAGCGCGGCGCTTCAGGCTGAATCTGTTCTTCCAGGGGAAACCACTTTATGACTCAA	120		
Qy	121	CAGCAGATGCTCTCTCTTTTCCGAGAAAGGATATTGATGACCCCTTTTGTCCTTTGAA	180		
Db	121	CAGCAGATGCTCTCTCTTTTCCGAGAAAGGATATTGATGACCCCTTTTGTCCTTTGAA	180		
Qy	181	GAAATGCAATGACCTCTCTGATGAAGATTAAAGCAGTGAAGCAACTTTGTCCGGAAGTAT	240		
Db	181	GAAATGCAATGACCTCTCTGATGAAGATTAAAGCAGTGAAGCAACTTTGTCCGGAAGTAT	240		

QY 241 TCCGACACCTAGCTAGTTACAGAGCTCCAGCCTTCGGCAAGACCTTCGAAGTCAGA 300
DB 241 TCCGACACCTAGCTAGTTACAGAGCTCCAGCCTTCGGCAAGACCTTCGAAGTCAGA 300
QY 301 AGCTTGTAGAGTGTGTGCTGCTTGTGAGAGTGAGAGTGTGTAAGAGAGCAACCGGG 360
DB 301 AGCTTGTAGAGTGTGTGCTGCTTGTGAGAGTGAGAGTGTGTAAGAGAGCAACCGGG 360
QY 361 GACATCATCTATGAAAGTGAAGAGAGAGCTTATTTGGCCAGAGAGAGTTTCA 420
DB 361 GACATCATCTATGAAAGTGAAGAGAGAGCTTATTTGGCCAGAGAGAGTTTCA 420
QY 421 TTTTGTGAGAGAGCGGACATTTATCTCGAAGCAAGCCCGTGGATCCCAATTA 480
DB 421 TTTTGTGAGAGAGCGGACATTTATCTCGAAGCAAGCCCGTGGATCCCAATTA 480
QY 481 CAGTATGCTCTTACAGAGCAAAATCACTTATCTGCTCATGGAATTCAGCCCTGAGGG 540
DB 481 CAGTATGCTCTTACAGAGCAAAATCACTTATCTGCTCATGGAATTCAGCCCTGAGGG 540
QY 541 GACTTGTCTCATCTTTGAAATAGATAGAGAGCAAGTTAGTAAACCTGATACAGTTT 600
DB 541 GACTTGTCTCATCTTTGAAATAGATAGAGAGCAAGTTAGTAAACCTGATACAGTTT 600
QY 601 TACCTAGCTAGCTGATTTTGGCTGTTACAGCGTTCACTGATGGAGTA GGTGCATCGA 660
DB 601 TACCTAGCTAGCTGATTTTGGCTGTTACAGCGTTCACTGATGGAGTA GGTGCATCGA 660
QY 661 GACATCAAGCTGAGAACTTTCTGTTGACCGACAGAGACATCAAGCTGGTGGATTTT 720
DB 661 GACATCAAGCTGAGAACTTTCTGTTGACCGACAGAGACATCAAGCTGGTGGATTTT 720
QY 721 GGATCTGCCGCAAAATGAATTCMAACAAGTGGTGAATGCCAACTCCGATTGGGACC 780
DB 721 GGATCTGCCGCAAAATGAATTCMAACAAGTGGTGAATGCCAACTCCGATTGGGACC 780
QY 781 CCAGATTACATGAGCTCTGAAAGTGTGACTGTGATGAAAGGAGATGAAAGGCACTTAC 840
DB 781 CCAGATTACATGAGCTCTGAAAGTGTGACTGTGATGAAAGGAGATGAAAGGCACTTAC 840
QY 841 GGGCTGACCTGTACCTGTGAGTGTGAGTGGGCTGATTTGCTTATGAGATTTATGGAGA 900
DB 841 GGGCTGACCTGTACCTGTGAGTGTGAGTGGGCTGATTTGCTTATGAGATTTATGGAGA 900
QY 901 TCCCTCTCGCAGAGGGAACCTCTGCGCAGAACCTTCAATTAACATTAATTTCCAGCG 960
DB 901 TCCCTCTCGCAGAGGGAACCTCTGCGCAGAACCTTCAATTAACATTAATTTCCAGCG 960
QY 961 TTTTGAATTTTCCAGATGACCCCAAGTGAAGAGAGTGAAGCTTTTGTGATCTGAATCAAGC 1020
DB 961 TTTTGAATTTTCCAGATGACCCCAAGTGAAGAGAGTGAAGCTTTTGTGATCTGAATCAAGC 1020
QY 1021 TTGTTGTGCGGCGCAAGAAAGAGAGCTGAAGTTGAAGGCTTTGCTGCCATCTTTTTC 1080
DB 1021 TTGTTGTGCGGCGCAAGAAAGAGAGCTGAAGTTGAAGGCTTTGCTGCCATCTTTTTC 1080
QY 1081 TCTAAATTTGATCTGGAACCAATTCGTAACCTCTCCCTCCCTGCTTCCACCTCAAG 1140
DB 1081 TCTAAATTTGATCTGGAACCAATTCGTAACCTCTCCCTCCCTGCTTCCACCTCAAG 1140
QY 1141 TCTGACGATGACCTTCAATTTTGAATGAACAGAGAGAAATTCGTGGTTTCACTCTCT 1200
DB 1141 TCTGACGATGACCTTCAATTTTGAATGAACAGAGAGAAATTCGTGGTTTCACTCTCT 1200
QY 1201 CCGTGCAGCTGAGCCCTCAGGCTTCTCGGCTGAAGAACTGCCGTTTGTGGGGTTTTCG 1260
DB 1201 CCGTGCAGCTGAGCCCTCAGGCTTCTCGGCTGAAGAACTGCCGTTTGTGGGGTTTTCG 1260
QY 1261 TACAGCAAGGCACTGGGGATTTCTGTAGATCTGAAGCTGTGTGCTGGGCTGGAGCTCC 1320
DB 1261 TACAGCAAGGCACTGGGGATTTCTGTAGATCTGAAGCTGTGTGCTGGGCTGGAGCTCC 1320
QY 1321 CTTGCCAAGACTAGCTTCATGAAAGAACTTCTCATCAAAAGCAAGAGCTACAGAC 1380

DB 1321 CTTGCCAAGACTAGCTTCATGAAAGAACTTCTCATCAAAAGCAAGAGCTACAGAC 1380
QY 1381 TCTCAGAGCAAGTGTACAAAGATGAGAGAGAAATGACCCGTTTACCTCGAGAGTCA 1440
DB 1381 TCTCAGAGCAAGTGTACAAAGATGAGAGAGAAATGACCCGTTTACCTCGAGAGTCA 1440
QY 1441 GAGTGAAGGCTGTCTTACTCAGAAAGAGTGAAGTGAAGGCTCTGAGACTCAGAGA 1500
DB 1441 GAGTGAAGGCTGTCTTACTCAGAAAGAGTGAAGTGAAGGCTCTGAGACTCAGAGA 1500
QY 1501 TCCCTCTGAGAGAGAGCTTGTCTTACTCAGAAAGTGAAGTGAAGGCTTAAAGCGAAGT 1560
DB 1501 TCCCTCTGAGAGAGAGCTTGTCTTACTCAGAAAGTGAAGTGAAGGCTTAAAGCGAAGT 1560
QY 1561 TTGAGCAAGCAGAGTGAAGTGTCTCCAGAGAGATGACAAACACTGACAGTTCTCAT 1620
DB 1561 TTGAGCAAGCAGAGTGAAGTGTCTCCAGAGAGATGACAAACACTGACAGTTCTCAT 1620
QY 1621 GATATCAGAGAGAGAGCCGGAAGCTTCAAGAAATCAAGAGAGAGTACAGGCTCAA 1680
DB 1621 GATATCAGAGAGAGAGCCGGAAGCTTCAAGAAATCAAGAGAGAGTACAGGCTCAA 1680
QY 1681 GTGGAAGAAATGAGTGTGATGAAATCACTTGGAAAGAGATCTTGTCTCAGCAAGAGA 1740
DB 1681 GTGGAAGAAATGAGTGTGATGAAATCACTTGGAAAGAGATCTTGTCTCAGCAAGAGA 1740
QY 1741 CGGAGTATCTTACGAATCTGAGCTGAGAGAGTCTGGCTTGTCTGCTGAAGATTCAG 1800
DB 1741 CGGAGTATCTTACGAATCTGAGCTGAGAGAGTCTGGCTTGTCTGCTGAAGATTCAG 1800
QY 1801 CGGAAAGCAGACAGATGTGACATTAACCTTTGAAGCTTAAGATCAAGGAAAGCTTGA 1860
DB 1801 CGGAAAGCAGACAGATGTGACATTAACCTTTGAAGCTTAAGATCAAGGAAAGCTTGA 1860
QY 1861 GTGGAAGATATCTCGAACTGGAAGAGATCAATGCTGAGAGAGCTCAAAATTCAGAG 1920
DB 1861 GTGGAAGATATCTCGAACTGGAAGAGATCAATGCTGAGAGAGAGCTCAAAATTCAGAG 1920
QY 1921 CTCGAAGAGAACTGGAAGAGCTGTAAAGCCAGAGCCGAGGCTGCTGCTCAG 1980
DB 1921 CTCGAAGAGAACTGGAAGAGCTGTAAAGCCAGAGCCGAGGCTGCTGCTCAG 1980
QY 1981 AATATCCGCGCAAGAGAGAGCCGAGAGAGAGCTGGAAGAGCTGCAAGCCGAGAG 2040
DB 1981 AATATCCGCGCAAGAGAGAGCCGAGAGAGAGCTGGAAGAGAGCTGCAAGCCGAGAG 2040
QY 2041 GATTTCTTGAAGGCAATCAGAAAGAGCTGTGAAGCTGAGAGAGCCGCCATTTCTCTG 2100
DB 2041 GATTTCTTGAAGGCAATCAGAAAGAGCTGTGAAGCTGAGAGAGCCGCCATTTCTCTG 2100
QY 2101 GAGAAGAGGTAAGAGACTGAGAGCAATGAGAGCTGAGAGAAACAGCTGAAGATGAC 2160
DB 2101 GAGAAGAGGTAAGAGACTGAGAGCAATGAGAGCTGAGAGAAACAGCTGAAGATGAC 2160
QY 2161 ATCCAGACAAATTCGAACAGATTCAGCAGATGGCTGATTAATTTCTGAGACTCGAAGAG 2220
DB 2161 ATCCAGACAAATTCGAACAGATTCAGCAGATGGCTGATTAATTTCTGAGACTCGAAGAG 2220
QY 2221 AAAATTCGAGAGGCGCAAGTCTCAGCCAGACCTTGAAGTGAACCTGAAGAGAG 2280
DB 2221 AAAATTCGAGAGGCGCAAGTCTCAGCCAGACCTTGAAGTGAACCTGAAGAGAG 2280
QY 2281 CAGCAGTGAAGAGAAAGATTAAGTGTGAGCAATCAGATTAAGAAAGCTGGCTGAC 2340
DB 2281 CAGCAGTGAAGAGAAAGATTAAGTGTGAGCAATCAGATTAAGAAAGCTGGCTGAC 2340
QY 2341 AAGAGACACTGAGAGACATGATCAGAGAGCAGAGAGAGGCGCTGAGAGAGGCGCAA 2400
DB 2341 AAGAGACACTGAGAGACATGATCAGAGAGCAGAGAGAGGCGCTGAGAGAGGCGCAA 2400
QY 2401 ATTCTAGCGAAGCAGAGGCGATGATCAATGCTATGAGATTCGAAGTCAAGATCCCTGAA 2460

Dh 2401 ATTCTACGCGAAGAGAGGCGATGATCAATGCTATGATTCGAAAGATCAAGATCCCTGGAA 2460
Qy 2461 CAGAGGATTTGGAACTGCTCTGAAGCCAAATTAACCTTGACGAAATAGAGATCTTTTAAAC 2520
Dh 2461 CAGAGGATTTGGAACTGCTCTGAAGCCAAATTAACCTTGACGAAATAGAGATCTTTTAAAC 2520
Qy 2521 CAAAGGAACATGAAGGCGCCAAAGAGATGATTTCTGAATCTCAGGCAACGAAATTTTAC 2580
Dh 2521 CAAAGGAACATGAAGGCGCCAAAGAGATGATTTCTGAATCTCAGGCAACGAAATTTTAC 2580
Qy 2581 CTGAGACACAGGCTGGGAAATTGGAGCCCAAGAACGAAACTGGAGAGCAGCTGGAG 2640
Dh 2581 CTGAGACACAGGCTGGGAAATTGGAGCCCAAGAACGAAACTGGAGAGCAGCTGGAG 2640
Qy 2641 AAGATCCGCCACCAAGACCAACAGTGAACAAGATCGCTGCTGGAACCTGGAGACAAGATTG 2700
Dh 2641 AAGATCCGCCACCAAGACCAACAGTGAACAAGATCGCTGCTGGAACCTGGAGACAAGATTG 2700
Qy 2701 GGGGAGGTCAGTCTAGAGCAAGAGGACAGAAATCGAGGCTCAAGCCCTGACAGAG 2760
Dh 2701 GGGGAGGTCAGTCTAGAGCAAGAGGACAGAAATCGAGGCTCAAGCCCTGACAGAG 2760
Qy 2761 CTACAGCTCTCCCTGACAGAGCGGAGTCAAGTTGACAGCCCTGACAGGCTGCAAGGCG 2820
Dh 2761 CTACAGCTCTCCCTGACAGAGCGGAGTCAAGTTGACAGCCCTGACAGGCTGCAAGGCG 2820
Qy 2821 GCCCTGAGAGCCAGCTTCCGCAAGCGAGACACAGAGCTGGAGAGACACAGCAGAAAGCT 2880
Dh 2821 GCCCTGAGAGCCAGCTTCCGCAAGCGAGACACAGAGCTGGAGAGACACAGCAGAAAGCT 2880
Qy 2881 GAAGAGAGATCCAGGCACTCAGGCGCATAGAGATGAATTCAGCGGCAAAATTTGAGGCT 2940
Dh 2881 GAAGAGAGATCCAGGCACTCAGGCGCATAGAGATGAATTCAGCGGCAAAATTTGAGGCT 2940
Qy 2941 CTTGCTAACAGCTGTACTGTATCAACAGACCTGGAGGAGGCTTAAACAGCTGACCGAG 3000
Dh 2941 CTTGCTAACAGCTGTACTGTATCAACAGACCTGGAGGAGGCTTAAACAGCTGACCGAG 3000
Qy 3001 GACACGCTGAATCTCAACCAAACTTCTACTTGTCAAACAATCGATGAGGCTTCT 3060
Dh 3001 GACACGCTGAATCTCAACCAAACTTCTACTTGTCAAACAATCGATGAGGCTTCT 3060
Qy 3061 GGGGCGCAAGAGATTGTAACTGCGAAGTGAAGTGAACATCTCCGCGCGGAGATC 3120
Dh 3061 GGGGCGCAAGAGATTGTAACTGCGAAGTGAAGTGAACATCTCCGCGCGGAGATC 3120
Qy 3121 ACGGAAACGAGATGAGCTTACCAACGAAACAAACGATGAGGCTTGAAGACCAAG 3180
Dh 3121 ACGGAAACGAGATGAGCTTACCAACGAAACAAACGATGAGGCTTGAAGACCAAG 3180
Qy 3181 TGCAACATGCTGAGGAACAAGTCAATGATTTGAGGCGCTTAAACGATGAGCTAGAA 3240
Dh 3181 TGCAACATGCTGAGGAACAAGTCAATGATTTGAGGCGCTTAAACGATGAGCTAGAA 3240
Qy 3241 AAAGAGCGCAGTGGAGGCGCTGGAGAGACCTCTCTGGGTGATGAGAAATCCAGTTTGAAG 3300
Dh 3241 AAAGAGCGCAGTGGAGGCGCTGGAGAGACCTCTCTGGGTGATGAGAAATCCAGTTTGAAG 3300
Qy 3301 TGTCCGGTTTCAAGAGCTGCAAGAGATGCTGACACCGAAGAAACAGAGACGGGAGAGGCG 3360
Dh 3301 TGTCCGGTTTCAAGAGCTGCAAGAGATGCTGACACCGAAGAAACAGAGACGGGAGAGGCG 3360
Qy 3361 GATCAGCGGATCACCGAGTCTCGCAGAGTGTGAGAGCTGGAGTGAAGGACCAAGGCT 3420
Dh 3361 GATCAGCGGATCACCGAGTCTCGCAGAGTGTGAGAGCTGGAGTGAAGGACCAAGGCT 3420
Qy 3421 GAGATTCTGCTCTGAGAGAGCTCTCAAAAGACAGAACTGAAGGCGAGAGCTCTCT 3480
Dh 3421 GAGATTCTGCTCTGAGAGAGCTCTCAAAAGACAGAACTGAAGGCGAGAGCTCTCT 3480
Qy 3481 GACAAAGCTCAATGACCTGAGAGAAAGCATGCTATCTTGAATGAAATGCCGAGAGCTTA 3540
Dh 3481 GACAAAGCTCAATGACCTGAGAGAAAGCATGCTATCTTGAATGAAATGCCGAGAGCTTA 3540

Qy 3541 CAGCAGAGCTGAGAGCTGAAACGAGAGCTCAACAGAGGCTTCTGAAAGCAAGCCAAA 3600
Dh 3541 CAGCAGAGCTGAGAGCTGAAACGAGAGCTCAACAGAGGCTTCTGAAAGCAAGCCAAA 3600
Qy 3601 TTACAGCAGAGATGAGCTGCGAGAAATCAATCTTTCCGTCTGATCTCAAGGCTGCA 3660
Dh 3601 TTACAGCAGAGATGAGCTGCGAGAAATCAATCTTTCCGTCTGATCTCAAGGCTGCA 3660
Qy 3661 GAAGCTCTAGATGAGGCTGATCTACTGAACAGAAAGAGTCACTTGAAGTACAGCTG 3720
Dh 3661 GAAGCTCTAGATGAGGCTGATCTACTGAACAGAAAGAGTCACTTGAAGTACAGCTG 3720
Qy 3721 GAAAACATTGAGGTTCTATATCTCATGAAAGGTGAAAATGAAAGGCACTATTTCTCAA 3780
Dh 3721 GAAAACATTGAGGTTCTATATCTCATGAAAGGTGAAAATGAAAGGCACTATTTCTCAA 3780
Qy 3781 CAAACCAAACTCATTTGATTTTCTGCAAGCCCAAAATGACCACTGCTTAAAAAGAAAG 3840
Dh 3781 CAAACCAAACTCATTTGATTTTCTGCAAGCCCAAAATGACCACTGCTTAAAAAGAAAG 3840
Qy 3841 GTTCCTCTGAGTACATGAGCTGAAGCTGGCCCTGGAGAAAGAAAGCTCGCTGCA 3900
Dh 3841 GTTCCTCTGAGTACATGAGCTGAAGCTGGCCCTGGAGAAAGAAAGCTCGCTGCA 3900
Qy 3901 GAGCTAGAGAAAGCCCTTCAAGAAACCCGATGAGCTCCGATCCGCGCGGAGAGAGCT 3960
Dh 3901 GAGCTAGAGAAAGCCCTTCAAGAAACCCGATGAGCTCCGATCCGCGCGGAGAGAGCT 3960
Qy 3961 GCCCAGCGCAAGCAACGACCAACCCCAACCCATCCACCGCAGCCACCGGAGAGCAGAG 4020
Dh 3961 GCCCAGCGCAAGCAACGACCAACCCCAACCCATCCACCGCAGCCACCGGAGAGCAGAG 4020
Qy 4021 ATGCCATGTCGCGCATCGTGCAGTGCAGAGCAACAGCCAGTCCATGAGCTGCTG 4080
Dh 4021 ATGCCATGTCGCGCATCGTGCAGTGCAGAGCAACAGCCAGTCCATGAGCTGCTG 4080
Qy 4081 GCCCGGCAATCCAGCCGCGAGAAAGAGTCTTCAACTCCAGAGAAATTTAGTGGCGTCTT 4140
Dh 4081 GCCCGGCAATCCAGCCGCGAGAAAGAGTCTTCAACTCCAGAGAAATTTAGTGGCGTCTT 4140
Qy 4141 AAGGAACGATGACCAACCAATTTCTCAACCCATTTCAAGTAGACTGAACATGCGAGCC 4200
Dh 4141 AAGGAACGATGACCAACCAATTTCTCAACCCATTTCAAGTAGACTGAACATGCGAGCC 4200
Qy 4201 ACAAAAGTGTGTGTGTGTGTGTGATACCTGACCTTGTGACGCGCAGGCTATCCAAATGCTC 4260
Dh 4201 ACAAAAGTGTGTGTGTGTGTGTGATACCTGACCTTGTGACGCGCAGGCTATCCAAATGCTC 4260
Qy 4261 GAAATGCAAGTATGTGTCAACCCAGTGTCTCAACGTTGCTTGCACGCACTTGCAGGCTTG 4320
Dh 4261 GAAATGCAAGTATGTGTCAACCCAGTGTCTCAACGTTGCTTGCACGCACTTGCAGGCTTG 4320
Qy 4321 CCTGCTGAATATGCGCACACACTTCAACGAGGCTTGTGCGGTGACAAATGAAATCTCCCA 4380
Dh 4321 CCTGCTGAATATGCGCACACACTTCAACGAGGCTTGTGCGGTGACAAATGAAATCTCCCA 4380
Qy 4381 GGTCTTCAGAACCAAGAGCCGACAGAGCTTGCACCTGGAAGGGGTGAATGAAGTGGCC 4440
Dh 4381 GGTCTTCAGAACCAAGAGCCGACAGAGCTTGCACCTGGAAGGGGTGAATGAAGTGGCC 4440
Qy 4441 AAGAAATTAACAAACGAGGACGCAAGGCTGGGACAGAAATGATCATTTGCTTGGAGGATCA 4500
Dh 4441 AAGAAATTAACAAACGAGGACGCAAGGCTGGGACAGAAATGATCATTTGCTTGGAGGATCA 4500
Qy 4501 AAAAGTCTCATTTATGACATTAAGACCAAGAAAGCTGGAACAAGGCGGTGGAAGAAATTT 4560
Dh 4501 AAAAGTCTCATTTATGACATTAAGAAACCAAGAAAGCTGGAACAAGGCGGTGGAAGAAATTT 4560
Qy 4561 GAGCTGTGCTTCCGAGCGGAGATGATATATTCATAGGAGCGGTGTGTTCTCCGAAGCTC 4620
Dh 4561 GAGCTGTGCTTCCGAGCGGAGATGATATATTCATAGGAGCGGTGTGTTCTCCGAAGCTC 4620

QY 4621 GCAAAATACAGCCAAAGAGATGTCTCCATACATCTGAAGATGGAATCTCAACCCGACACCC 4680
 Db 4621 GCAAAATACAGCCAAAGAGATGTCTCCATACATCTGAAGATGGAATCTCAACCCGACACCC 4680
 QY 4681 ACCGCTGAGCCGCGGAGAACCTCTACTGTGTAGCTCCGAGCTTCCGTCACAAACAGGGC 4740
 Db 4681 ACCGCTGAGCCGCGGAGAACCTCTACTGTGTAGCTCCGAGCTTCCGTCACAAACAGGGC 4740
 QY 4741 TGGGTCAACCGCTTAGATCAGTGTGTGAGGAGAGTGTCTAGGAGAAAGACAGAA 4800
 Db 4741 TGGGTCAACCGCTTAGATCAGTGTGTGAGGAGAGTGTCTAGGAGAAAGACAGAA 4800
 QY 4801 GCTGATCTTAACTGCTTTGGAATCTCCCTGTGAACTCGAAGGTGATGACCGCTTAC 4860
 Db 4801 GCTGATCTTAACTGCTTTGGAATCTCCCTGTGAACTCGAAGGTGATGACCGCTTAC 4860
 QY 4861 ATGAACCTGACGCGCTTCACTGACGAGTGTGTGAGGAGACCGAGGAGAGGAGCTC 4920
 Db 4861 ATGAACCTGACGCGCTTCACTGACGAGTGTGTGAGGAGACCGAGGAGAGGAGCTC 4920
 QY 4921 TAGCCCTGATGCTTGAAGAACTCCCTTAAACCATGTCACGGAATTGAGAGAGCTTC 4980
 Db 4921 TAGCCCTGATGCTTGAAGAACTCCCTTAAACCATGTCACGGAATTGAGAGAGCTTC 4980
 QY 4981 CAAATTTATATATCAAGAGCTTGAGAGAGCTCATGATGAGAGAGAGAGAGAGGAGCA 5040
 Db 4981 CAAATTTATATATCAAGAGAGCTTGAGAGAGCTCATGATGAGAGAGAGAGAGGAGCA 5040
 QY 5041 CTGTGCTTGTGAGAGTGAAGAAAGTGAACAGTCCCTGAGGAGCTCCAGCTGAGCTGCC 5100
 Db 5041 CTGTGCTTGTGAGAGTGAAGAAAGTGAACAGTCCCTGAGGAGCTCCAGCTGAGCTGCC 5100
 QY 5101 CAGCCGACATCTCACCACCAATTTTGAAGCTGCAAGGAGCTCCAGCTTGTGAGGAGCA 5160
 Db 5101 CAGCCGACATCTCACCACCAATTTTGAAGCTGCAAGGAGCTCCAGCTTGTGAGGAGCA 5160
 QY 5161 GGCAGATGTAAGAGAGGAGCTGTGATGTGACAGCCATGCCAGCAAAAGTGTCTCTC 5220
 Db 5161 GGCAGATGTAAGAGAGGAGCTGTGATGTGACAGCCATGCCAGCAAAAGTGTCTCTC 5220
 QY 5221 CGCTACACGAAAACTTCAGCAATCTGCAATCCGGAAGAGATGAGAGCTTCAGAGCCC 5280
 Db 5221 CGCTACACGAAAACTTCAGCAATCTGCAATCCGGAAGAGATGAGAGCTTCAGAGCCC 5280
 QY 5281 TGCAGCTGTATCCACTTCAACCAATTAAGATCTCTCATTTGGAACCAATTAATTTAGGAA 5340
 Db 5281 TGCAGCTGTATCCACTTCAACCAATTAAGATCTCTCATTTGGAACCAATTAATTTAGGAA 5340
 QY 5341 ATGCACATGAGAGAGTACAGGCTTCAGAGAAATCTGTGATGAGATGAGACATTCCTGGCA 5400
 Db 5341 ATGCACATGAGAGAGTACAGGCTTCAGAGAAATCTGTGATGAGATGAGACATTCCTGGCA 5400
 QY 5401 CTGTGCTGTGTGCTGCTCTTCCAGAGCTTCCCTGTCTCAATGTGACAGGTGACAGC 5460
 Db 5401 CTGTGCTGTGTGCTGCTCTTCCAGAGCTTCCCTGTCTCAATGTGACAGGTGACAGC 5460
 QY 5461 GCAAGGAGAGAGAGAGTACTTGTGTGTGAGGAGTGTGAGAGTGTGAGAGTGTCT 5520
 Db 5461 GCAAGGAGAGAGAGAGTACTTGTGTGTGAGGAGTGTGAGAGTGTGAGAGTGTCT 5520
 QY 5521 TACGAGAGAGTACCGCAGACGATCTCAAGTGAAGTGCCTTACCTTGGCCTTGGCC 5580
 Db 5521 TACGAGAGAGTACCGCAGACGATCTCAAGTGAAGTGCCTTACCTTGGCCTTGGCC 5580
 QY 5581 TACAGAGAACCTATCTGTTTGTGACCACTTCACTCACTCACTCACTCACTCACTCACTCACT 5640
 Db 5581 TACAGAGAACCTATCTGTTTGTGACCACTTCACTCACTCACTCACTCACTCACTCACTCACT 5640
 QY 5641 GCAAGCTCTCAG 5700
 Db 5641 GCAAGCTCTCAG 5700
 QY 5701 CTGGGCTCTGCAATTCCTCAGAGAGATTTATCTTGGAGTCTCTCATACAGAGATTAATTA 5760

Db 5701 CTGGGCTCTGCAATTCCTCAGAGAGATTTATCTTGGAGTCTCTCATACAGAGATTAATTA 5760
 QY 5761 AGGGTCAATTTGCTGCAAGAGAGAACTTGTGAAGAGTCCGAGCACTGAACCAACCGGAGC 5820
 Db 5761 AGGGTCAATTTGCTGCAAGAGAGAACTTGTGAAGAGTCCGAGCACTGAACCAACCGGAGC 5820
 QY 5821 CCGTCCACTCCGAG 5880
 Db 5821 CCGTCCACTCCGAG 5880
 QY 5881 ACCAAGCGCTGAGCTTCCAG 5940
 Db 5881 ACCAAGCGCTGAGCTTCCAG 5940
 QY 5941 AGCAGACCCCAACGCTTCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6000
 Db 5941 AGCAGACCCCAACGCTTCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6000
 QY 6001 CCGCCCTGAG 6060
 Db 6001 CCGCCCTGAG 6060
 QY 6061 CCGGAGAGAGCTGTTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6120
 Db 6061 CCGGAGAGAGCTGTTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6120
 QY 6121 CCGCTGCTCCAGGTGAACAGAGTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6165
 Db 6121 CCGCTGCTCCAGGTGAACAGAGTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6165
 RESULT 3
 AX504254 6298 bp DNA linear PAT 27-SEP-2002
 LOCUS AX504254
 DEFINITION Sequence 43 from Patent WO0233099.
 ACCESSION AX504254
 VERSION AX504254.1 GI:23386094
 KEYWORDS
 SOURCE
 ORGANISM
 Homo sapiens (human)
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE
 1
 AUTHORS Gururajan, R., Baughn, M.R., Walla, N.K., Elliott, V.S., Xu, Y., Arvizu, C., Yao, M.G., Ramkumar, J., Ding, L., Tang, Y.T., Hafalia, A.J., Nguyen, D.B., Gandhi, A.R., Lu, Y., Yee, H., Burford, N., Bandman, O., Tribouley, C.M., Lal, P.G., Recipon, S.A., Lu, D.A., Botowsky, M.L., Thornton, M., Swarnaker, A., Thangavelu, K., Khan, F.A. and Ison, C.H.
 TITLE Human kinases
 JOURNAL Patent: WO 0233099-A 43 25-APR-2002;
 Incyte Genomics, Inc. (US)
 FEATURES
 source
 1..6298
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"
 /note="Incyte ID No: 7484498CB1"
 ORIGIN
 Query Match 99.9%; Score 6161.8; DB 6; Length 6298;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 6163; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 ATGTTGAAGTTCAAAATATGAGAGCGGGAATCTTTGATGCTGTGCTGAACCAATT 60
 Db 55 ATGTTGAAGTTCAAAATATGAGAGCGGGAATCTTTGATGCTGTGCTGAACCAATT 114
 QY 61 GCCAGCGGAGCTTCAG 120
 Db 115 GCCAGCGGAGCTTCAG 174
 QY 121 CAGCAGATGTCTCTCTTTTCCGAGAGAGAGATTAAGATGCCCTTGTGTTCTTTGAA 180

Db 175 CAGAGATGTTCTCTTTCCCGAAGGATTAATGATGCCCTCTTTGTTCTTTGAA 234
Qy 181 GAATGACGTAGCCTGCTCTGATGAAGATTAAAGCAGTGAACAATTGTCCGGAATAT 240
Db 235 GAATGACGTAGCCTGCTCTGATGAAGATTAAAGCAGTGAACAATTGTCCGGAATAT 294
Qy 241 TCCGACACCAATAGCTGATTACAGAGCTCCAGCTTCCGAAAGACCTTGAAGTCA 300
Db 295 TCCGACACCAATAGCTGATTACAGAGCTCCAGCTTCCGAAAGACCTTGAAGTCA 354
Qy 301 AGCTTGTAAGTTGTTGTCATCTTGTGTAAGTGCAGGTGTAAAGAGAAACAACCGG 360
Db 355 AGCTTGTAAGTTGTTGTTGTCATCTTGTGTAAGTGCAGGTGTAAAGAGAAACAACCGG 414
Qy 361 GACATCTATGCTATGAAGATGATGAAGAGAGGCTTATTTGCCCAGAGCAGTTTCA 420
Db 415 GACATCTATGCTATGAAGATGATGAAGAGAGGCTTATTTGCCCAGAGCAGTTTCA 474
Qy 421 TTTTTGAGAGAGCCGGAACATTTATCTCGAAGCAACAGCCGTGATCCCAATTGA 480
Db 475 TTTTTGAGAGAGCCGGAACATTTATCTCGAAGCAACAGCCGTGATCCCAATTGA 534
Qy 481 CAGTATGCTTTGAGACAAATAATCACCTTATCTGGTCATGGAATATCAGCCTGAGGG 540
Db 535 CAGTATGCTTTGAGACAAATAATCACCTTATCTGGTCATGGAATATCAGCCTGAGGG 594
Qy 541 GACTTGCTGTCATTTTGAATAGATATGAGACCAAGTTAGATGAACCTGATACAGTTT 600
Db 595 GACTTGCTGTCATTTTGAATAGATATGAGACCAAGTTAGATGAACCTGATACAGTTT 654
Qy 601 TACCTAGCTGAGCTGATTTTGGCTGTTCAACAGCTTATCTGATGGGATACGTGATGA 660
Db 655 TACCTAGCTGAGCTGATTTTGGCTGTTCAACAGCTTATCTGATGGGATACGTGATGA 714
Qy 661 GACATCAAGGCTGAGAACATTTCTGTTGACCGCACAGAGACATCAAGCTGGTGAATTT 720
Db 715 GACATCAAGGCTGAGAACATTTCTGTTGACCGCACAGAGACATCAAGCTGGTGAATTT 774
Qy 721 GGATCTGCCGCGAATAATGAATTCAACAGATGGTGAATGCCAACTCCGATTTGGAGC 780
Db 775 GGATCTGCCGCGAATAATGAATTCAACAGATGGTGAATGCCAACTCCGATTTGGAGC 834
Qy 781 CCAGATTAATGAGCTCTGTAAGTCTGATCTGTATGAACGGGATGAGAAAGCACTTAC 840
Db 835 CCAGATTAATGAGCTCTGTAAGTCTGATCTGTATGAACGGGATGAGAAAGCACTTAC 894
Qy 841 GGCTGGAATGTAAGTGTGATGAGTGGGCGTGAATGGCTTATGAGATGATTAATGGAGA 900
Db 895 GGCTGGAATGTAAGTGTGATGAGTGGGCGTGAATGGCTTATGAGATGATTAATGGAGA 954
Qy 901 TCCCTCTGCGAGAGGAACTCTGCGACAGACCTTCAATTAATGAATTTCCAGCG 960
Db 955 TCCCTCTGCGAGAGGAACTCTGCGACAGACCTTCAATTAATGAATTTCCAGCG 1014
Qy 961 TTTTGAATTTTCCATATGACCCCAAGTGAAGTGAATTTTGAATCTGATCTGAATCAAGC 1020
Db 1015 TTTTGAATTTTCCATATGACCCCAAGTGAAGTGAATTTTGAATCTGATCTGAATCAAGC 1074
Qy 1021 TTGTTGTGGGCGAGAAAGAGACTGAAGTTGAAGTCTTGTGCGACATCCTTTCTTC 1080
Db 1075 TTGTTGTGGGCGAGAAAGAGACTGAAGTTGAAGTCTTGTGCGACATCCTTTCTTC 1134
Qy 1081 TCTAAATTTGACTGGAACAACATTCGTAATCTCTCTCCCTTCGTTCCACCTCAAG 1140
Db 1135 TCTAAATTTGACTGGAACAACATTCGTAATCTCTCTCCCTTCGTTCCACCTCAAG 1194
Qy 1141 TCTGACGATGACCTTCAATTTTGAATGAACAGAGAAAGATTCGTGGGTTTCACTCTT 1200
Db 1195 TCTGACGATGACCTTCAATTTTGAATGAACAGAGAAAGATTCGTGGGTTTCACTCTT 1254
Qy 1201 CGGTGCACTGAGCCCTCAGGCTTCTGGGTGAAGAACTGCCGTTTGTGGGTTTTCG 1260

Db 1255 CCGTCCAGCTGAGCCCTCAGGCTTCTCGGTTGAAGAACTGCCGTTTGTGGGTTTTCG 1314
Qy 1261 TACAGCAAGGCACTGGGAAATCTTGTGTAGATCTGAGTCTGTGTTGTGCGGCTGCACTCC 1320
Db 1315 TACAGCAAGGCACTGGGAAATCTTGTGTAGATCTGAGTCTGTGTTGTGCGGCTGCACTCC 1374
Qy 1321 CTTGCCAAGCTAGCTCATTGGAAGAACTTCTCATCAAAAGCAAGACTCAAGAC 1380
Db 1375 CTTGCCAAGCTAGCTCATTGGAAGAACTTCTCATCAAAAGCAAGACTCAAGAC 1434
Qy 1381 TCTCAGCAAGTGTCACAAGATGAGCAGAAATGACCCGGTTTACTCCGAGGTGTCA 1440
Db 1435 TCTCAGCAAGTGTCACAAGATGAGCAGAAATGACCCGGTTTACTCCGAGGTGTCA 1494
Qy 1441 GAGTGAAGGCTGTGCTTACTGAAAGAGTGAAGTGAAGGCTTGAAGCTCAGAGA 1500
Db 1495 GAGTGAAGGCTGTGCTTACTGAAAGAGTGAAGTGAAGGCTTGAAGCTCAGAGA 1554
Qy 1501 TCCCTCTGAGCAGAGACTTGTCTACTACATCAAGAAATGCAAGTACTTAAAGCGAAT 1560
Db 1555 TCCCTCTGAGCAGAGACTTGTCTACTACATCAAGAAATGCAAGTACTTAAAGCGAAT 1614
Qy 1561 TTGAGCAAGCAGAGATGAGAGTGTCCAGAGAGATGACAAAGCACTGACGTTTCCAT 1620
Db 1615 TTGAGCAAGCAGAGATGAGAGTGTCCAGAGAGATGACAAAGCACTGACGTTTCCAT 1674
Qy 1621 GATATCAGAGCAGAGCCCGAAGCTCAAGAAATCAAGAGCAGAGTACAGGCTCAA 1680
Db 1675 GATATCAGAGCAGAGCCCGAAGCTCAAGAAATCAAGAGCAGAGTACAGGCTCAA 1734
Qy 1681 GTGGAAGAAATGAGTTGATGAAATCAGTTGGAAGAGATCTTGTCTCAGCAAGAGA 1740
Db 1735 GTGGAAGAAATGAGTTGATGAAATCAGTTGGAAGAGATCTTGTCTCAGCAAGAGA 1794
Qy 1741 CGGAGTATCTCTAAGAACTGTGAGTGTGAGAGTCTCGGCTTCTGCTGAAGAAATTCAG 1800
Db 1795 CGGAGTATCTCTAAGAACTGTGAGTGTGAGAGTCTCGGCTTCTGCTGAAGAAATTCAG 1854
Qy 1801 CGGAAAGCAGCAAGATGTGAGCAATAACTGTTGAAGCTTAAGATCAAGGAAAGCCTGAA 1860
Db 1855 CGGAAAGCAGCAAGATGTGAGCAATAACTGTTGAAGCTTAAGATCAAGGAAAGCCTGAA 1914
Qy 1861 GTGGAGAAATATGCCAACTGGAAAGATCATGCTGACAGCAGCTCAAAATTCAGAG 1920
Db 1915 GTGGAGAAATATGCCAACTGGAAAGATCATGCTGACAGCAGCTCAAAATTCAGAG 1974
Qy 1921 CTCCAAGAGAACTGGAGAAAGCTGTAAAGCCAGACGGAAGCCAGAGCTGCTGAC 1980
Db 1975 CTCCAAGAGAACTGGAGAAAGCTGTAAAGCCAGACGGAAGCCAGAGCTGCTGAC 2034
Qy 1981 AATATCCGCAAGCAAGAGCGAGCCGAGAGCTGAGAGTGTGAGAACCTGAGAG 2040
Db 2035 AATATCCGCAAGCAAGAGCGAGCCGAGAGCTGAGAGTGTGAGAACCTGAGAG 2094
Qy 2041 GATTCTTCTGAAAGCATCAAGAAAGAGCTGTGGAAGCTGAGAAAGCCGCCATTTCTTG 2100
Db 2095 GATTCTTCTGAAAGCATCAAGAAAGAGCTGTGGAAGCTGAGAAAGCCGCCATTTCTTG 2154
Qy 2101 GAGAACAAAGTAAAGACTAAGAACCAATGAGCCGTAAGAAACAGACTGAAGATGAC 2160
Db 2155 GAGAACAAAGTAAAGACTAAGAACCAATGAGCCGTAAGAAACAGACTGAAGATGAC 2214
Qy 2161 ATCCAGCAAAATCCCAACAGATCCAGCAGATGAGCTGATTAATTTCTGAGCTCGAAGAG 2220
Db 2215 ATCCAGCAAAATCCCAACAGATCCAGCAGATGAGCTGATTAATTTCTGAGCTCGAAGAG 2274
Qy 2221 AAACATCGGAGGCCCAAGTCTCAGCCCAAGCACTTAAGAGTCACTTGAACAGAAAGAG 2280
Db 2275 AAACATCGGAGGCCCAAGTCTCAGCCCAAGCACTTAAGAGTCACTTGAACAGAAAGAG 2334
Qy 2281 CAGCACTATGAGAGAAAGATTAAAGTTTGAACAATTCAGATTAAGAAAGACCTGTGCTGAC 2340
Db 2335 CAGCACTATGAGAGAAAGATTAAAGTTTGAACAATTCAGATTAAGAAAGACCTGTGCTGAC 2394

QY 2341 AAGGAGACTGGAAATGATGTCAGAGACAGAGAGAGGCCCTGAGAAAGGCCAAA 2400
| | | | |
Db 2395 AAGGAGACACTGGAGAACTGATGTCAGAGACAGAGAGAGGCCCTGAGAAAGGCCAAA 2454
| | | | |
QY 2401 ATTCTGCGCAACGAGAGGCGATGATCAATGCTATGGAATTCAGAGATCAGATCCCTGAA 2460
| | | | |
Db 2455 ATTCTGCGCAACGAGAGGCGATGATCAATGCTATGGAATTCAGAGATCAGATCCCTGAA 2514
| | | | |
QY 2461 CAGAGATTGTGAACTGTCTGAAGCCAAATTAATTGACCAATATGACAGTCTTTTACC 2520
| | | | |
Db 2515 CAGAGATTGTGAACTGTCTGAAGCCAAATTAATTGACCAATATGACAGTCTTTTACC 2574
| | | | |
QY 2521 CAAAGGAACTGAAGGCCCAAGAAAGATGATTTCTGAATCTCAGGCCAAGAAATTTTAC 2580
| | | | |
Db 2575 CAAAGGAACTGAAGGCCCAAGAAAGATGATTTCTGAATCTCAGGCCAAGAAATTTTAC 2634
| | | | |
QY 2581 CTGGAGACAGAGGCTGGGAATTTGAGGCCAGAACCCGAAAATGGAGAGCGAGCTGGAG 2640
| | | | |
Db 2635 CTGGAGACAGAGGCTGGGAATTTGAGGCCAGAACCCGAAAATGGAGAGCGAGCTGGAG 2694
| | | | |
QY 2641 AAGATCAGCCACCAAGACCAAGTGAACAAGATCGGCTGTGAACTGGAGACAAGATTG 2700
| | | | |
Db 2695 AAGATCAGCCACCAAGACCAAGTGAACAAGATCGGCTGTGAACTGGAGACAAGATTG 2754
| | | | |
QY 2701 CGGAGGTCAGTCTAGAGCAAGAGAGCAGAAACTGAGCTCAAGCCGACGCTCACAGAG 2760
| | | | |
Db 2755 CGGAGGTCAGTCTAGAGCAAGAGAGCAGAAACTGAGGCTCAAGCCGACGCTCACAGAG 2814
| | | | |
QY 2761 CTACAGCTCTCCCTGCGAGAGCGGAGTCAAGTTGACAGCCCTGCGAGCTGACGGGCG 2820
| | | | |
Db 2815 CTACAGCTCTCCCTGCGAGAGCGGAGTCAAGTTGACAGCCCTGCGAGCTGACGGGCG 2874
| | | | |
QY 2821 GCCCTGAGAGCCGAGCTTCGCGAGCGAAGACAGAGCTGGAGAGACACAGCGAAAGCT 2880
| | | | |
Db 2875 GCCCTGAGAGCCGAGCTTCGCGAGCGAAGACAGAGCTGGAGAGACACAGCGAAAGCT 2934
| | | | |
QY 2881 GAAAGAGAGATCCAGGCACTCAGCGCAATAGAGATGAATCCAGCCAAATTTGATGCT 2940
| | | | |
Db 2935 GAAAGAGAGATCCAGGCACTCAGCGCAATAGAGATGAATCCAGCCAAATTTGATGCT 2994
| | | | |
QY 2941 CTTCGTAAACAGCTGTATCTGTATATACAGACTTGGAGAGACGCTAAACAGCTGACCGAG 3000
| | | | |
Db 2995 CTTCGTAAACAGCTGTATCTGTATATACAGACTTGGAGAGACGCTAAACAGCTGACCGAG 3054
| | | | |
QY 3001 GACAAGCTGAACCTCAACCAACCAAACTTCTACTGTCCAAACCACTCGATGAGGCTTCT 3060
| | | | |
Db 3055 GACAAGCTGAACCTCAACCAACCAAACTTCTACTGTCCAAACCACTCGATGAGGCTTCT 3114
| | | | |
QY 3061 GGGGCCAAGCAGAGATTTGTACAACTGCAAGTGAAGTGCACATCTCCGCGGAGATC 3120
| | | | |
Db 3115 GGGGCCAAGCAGAGATTTGTACAACTGCAAGTGAAGTGCACATCTCCGCGGAGATC 3174
| | | | |
QY 3121 ACGGAACGAGAGATGACGCTTACAGCCAGAAAGCAACGATGAGGCTTGAAGACCAAG 3180
| | | | |
Db 3175 ACGGAACGAGAGATGACGCTTACAGCCAGAAAGCAACGATGAGGCTTGAAGACCAAG 3234
| | | | |
QY 3181 TGCACCACTGCTGAGAGAAACAGGTATATGATTTGAGAGCCCTTAACGATGAGCTGTGAA 3240
| | | | |
Db 3235 TGCACCACTGCTGAGAGAAACAGGTATATGATTTGAGAGCCCTTAACGATGAGCTGTGAA 3294
| | | | |
QY 3241 AAAGAGCGGAGTGGAGAGCCCTGAGAGAGCGTCTGGGTGATGAGAAATCCCAATTTGAG 3300
| | | | |
Db 3295 AAAGAGCGGAGTGGAGAGCCCTGAGAGAGCGTCTGGGTGATGAGAAATCCCAATTTGAG 3354
| | | | |
QY 3301 TGTCTGGATTGAGAGCTGCGAGAGATGCTGGAACAACGAAACAGACAGGCGAGAGCC 3360
| | | | |
Db 3355 TGTCTGGATTGAGAGCTGCGAGAGATGCTGGAACAACGAAACAGACAGGCGAGAGCC 3414
| | | | |
QY 3361 GATTAAGGAGATCAACGAGTCTCCGACAGTGTGAGAGCTGCGACGTGAAGAGCAAAAGGCT 3420
| | | | |
Db 3415 GATTAAGGAGATCAACGAGTCTCCGACAGTGTGAGAGCTGCGACGTGAAGAGCAAAAGGCT 3474
| | | | |

QY 3421 GAGATTCTGCTCTGACAGAGGCTCTCAAGACAGAAAGCTGAAGGCCGAGAGCTCTCT 3480
| | | | |
Db 3475 GAGATTCTGCTCTGACAGAGGCTCTCAAGACAGAAAGCTGAAGGCCGAGAGCTCTCT 3534
| | | | |
QY 3481 GACAGCTCAATGACCTGAGAAAGACATGCTATGCTTGAATTAATGAAATGCCGGAAGCTTA 3540
| | | | |
Db 3535 GACAGCTCAATGACCTGAGAAAGACATGCTATGCTTGAATTAATGAAATGCCGGAAGCTTA 3594
| | | | |
QY 3541 CAGCAGAGCTGAGACCTGAACGAGAGCTCAAAACAGAGGCTTTGGAAGACCAAGCCAAA 3600
| | | | |
Db 3595 CAGCAGAGCTGAGACCTGAACGAGAGCTCAAAACAGAGGCTTTGGAAGACCAAGCCAAA 3654
| | | | |
QY 3601 TTACAGTACAGATGAGACCTGCGAAGAAAATCAATTTCCGCTGATCAAGACCTGCAA 3660
| | | | |
Db 3655 TTACAGTACAGATGAGACCTGCGAAGAAAATCAATTTCCGCTGATCAAGACCTGCAA 3714
| | | | |
QY 3661 GAAGCTCTAGATGGGCTGATCTACTGAAGACAGAAAGATGACTTGGAGTATCAGCTG 3720
| | | | |
Db 3715 GAAGCTCTAGATGGGCTGATCTACTGAAGACAGAAAGATGACTTGGAGTATCAGCTG 3774
| | | | |
QY 3721 GAAACATTCAGGTTCTCTATTTCTCATGAAAAGTGAATAATGAAAGCACTATTTCTCAA 3780
| | | | |
Db 3775 GAAACATTCAGGTTCTCTATTTCTCATGAAAAGTGAATAATGAAAGCACTATTTCTCAA 3834
| | | | |
QY 3781 CAAACCAACTCATTTGATTTTCTGCAAGCCAAAATGSAACCACTGCTTAATAAAGAAAAG 3840
| | | | |
Db 3835 CAAACCAACTCATTTGATTTTCTGCAAGCCAAAATGSAACCACTGCTTAATAAAGAAAAG 3894
| | | | |
QY 3841 GTTCCTGCAAGTACAAATGAGCTGAAGCTGGCCCTGSAAGAAAGTAAAGCTCGCTGCA 3900
| | | | |
Db 3895 GTTCCTGCAAGTACAAATGAGCTGAAGCTGGCCCTGSAAGAAAGTAAAGCTCGCTGCA 3954
| | | | |
QY 3901 GAGCTAGAGAAACCCCTTCAAGAAAGCCGAGTGAAGCTCCGCTCCGCGSAGAAAGCT 3960
| | | | |
Db 3955 GAGCTAGAGAAACCCCTTCAAGAAAGCCGAGTGAAGCTCCGCTCCGCGSAGAAAGCT 4014
| | | | |
QY 3961 GCCCACCAGCAAGCAAGACCAACCAACCTTCAACGCGACCAACCGAGCAGCAGCAG 4020
| | | | |
Db 4015 GCCCACCAGCAAGCAAGACCAACCAACCTTCAACGCGACCAACCGAGCAGCAGCAG 4074
| | | | |
QY 4021 ATGCGCATGTCGCCCATGCGGCGGTGCGCAGAGACCAAGCCAGTGCATGAGCTGCTG 4080
| | | | |
Db 4075 ATGCGCATGTCGCCCATGCGGCGGTGCGCAGAGACCAAGCCAGTGCATGAGCTGCTG 4134
| | | | |
QY 4081 GCCCGGCATCCAGCCGCAAGAAAGAGTCTTCAATCTCAGAGAAATTTAGTCCGCTCTT 4140
| | | | |
Db 4135 GCCCGGCATCCAGCCGCAAGAAAGAGTCTTCAATCTCAGAGAAATTTAGTCCGCTCTT 4194
| | | | |
QY 4141 AAGGAACGATGACCAACAATATTTCTCACCGATTCAACGTAGAGCTGAACATGCGAGCC 4200
| | | | |
Db 4195 AAGGAACGATGACCAACAATATTTCTCACCGATTCAACGTAGAGCTGAACATGCGAGCC 4254
| | | | |
QY 4201 ACAAAAGTGTCTGTGTCTGTGATACCGTGCATTTTGAAGCCAGGCAATCCAAATGTCTC 4260
| | | | |
Db 4255 ACAAAAGTGTCTGTGTCTGTGATACCGTGCATTTTGAAGCCAGGCAATCCAAATGTCTC 4314
| | | | |
QY 4261 GAATGTCAGTGAATGTGTCACCCCAAGTGTCCACGTCGCTTGCACAGCACTGCGGCTTG 4320
| | | | |
Db 4315 GAATGTCAGTGAATGTGTCACCCCAAGTGTCTCCACGTCGCTTGCACAGCACTGCGGCTTG 4374
| | | | |
QY 4321 CTTGCTGAATATATGACACACACTTCAACGAGGCTTCTGCGGTGACAAATGAATCTCCCA 4380
| | | | |
Db 4375 CTTGCTGAATATATGACACACACTTCAACGAGGCTTCTGCGGTGACAAATGAATCTCCCA 4434
| | | | |
QY 4381 GGTCTCAGACCAAGAGCCCAAGACGAGCTTGAACCTGGAAGGCTGATGAAAGTTCCTC 4440
| | | | |
Db 4435 GGTCTCAGACCAAGAGCCCAAGACGAGCTTGAACCTGGAAGGCTGATGAAAGTTCCTC 4494
| | | | |
QY 4441 AAGGAATTAACAAAGAGACAGCAAGGCTGGAGACAGAAATGATATGCTCTGAGAGATCA 4500
| | | | |
Db 4495 AAGGAATTAACAAAGAGACAGCAAGGCTGGAGACAGAAATGATATGCTCTGAGAGATCA 4554
| | | | |
QY 4501 AAGTCTCTATTTATGACAAATGAAGCCAGAAAGCTGGACAGAGGCCGCTGGAAGAAATTT 4560
| | | | |


```

Db      4555 AAGTCCCTATTATGACAAATGAAAGCCAGAAAGCTGGACGAGAGCCGGTGGAAATTT 4614
Qy      4561 GAGGTGCTCTTCCCGACGGGGATGTATCTATTCATGCTGCGTGTGCTTCCGAATC 4620
Db      4615 GAGGTGCTCTTCCCGACGGGGATGTATCTATTCATGCTGCGTGTGCTTCCGAATC 4674
Qy      4621 GCAATATACAGCAAGAGAGATGTCCCATACATCTAGTAAGATGTATCTACCCGACACC 4680
Db      4675 GCAATATACAGCAAGAGATGTCCCATACATCTAGTAAGATGTATCTACCCGACACC 4734
Qy      4681 ACCTGTGCGCCGGGAGAACCTCTACTTGTAGCTCCAGCTTCCCTGACAAACAGGCG 4740
Db      4735 ACCTGTGCGCCGGGAGAACCTCTACTTGTAGCTCCAGCTTCCCTGACAAACAGGCG 4794
Qy      4741 TGGGTCAACCGCTTGAATCATGTTGTGCGAGGTGGAGAGTTTCTAGGAAAAAGCAAA 4800
Db      4795 TGGGTCAACCGCTTGAATCATGTTGTGCGAGGTGGAGAGTTTCTAGGAAAAAGCAAA 4854
Qy      4801 GGTGATGCTAAACGCTTGGAAAACTCCCTGTGTAACCTGGAAAGGTGATGACCTGTAGAC 4860
Db      4855 GGTGATGCTAAACGCTTGGAAAACTCCCTGTGTAACCTGGAAAGGTGATGACCTGTAGAC 4914
Qy      4861 ATGAATGACAGCTGACCTTCACTGACCAAGTGTGTGTGGGACCGAGAGAGGCTC 4920
Db      4915 ATGAATGACAGCTGACCTTCACTGACCAAGTGTGTGTGGGACCGAGAGAGGCTC 4974
Qy      4921 TACGCTGATGTCTTGAAGAACTCCCTAACCCATGTCCAGGAATTGGAGAGTCTTC 4980
Db      4975 TACGCTGATGTCTTGAAGAACTCCCTAACCCATGTCCAGGAATTGGAGAGTCTTC 5034
Qy      4981 CAAATTTATTTATCAAGAGACTGGAGAACTCTCATGTATACAGAGAAAGCGGGCA 5040
Db      5035 CAAATTTATTTATCAAGAGACTGGAGAACTCTCATGTATACAGAGAAAGCGGGCA 5094
Qy      5041 CTGTGTCTTGTGACGTGAAGAAAGTGAACAGTCCCTGCGCCAGTCCCACTGCTGCC 5100
Db      5095 CTGTGTCTTGTGACGTGAAGAAAGTGAACAGTCCCTGCGCCAGTCCCACTGCTGCC 5154
Qy      5101 CAGCCCGACATCTACCCAAATTTTGAAGCTGTCAAGGGCTGCCACTTGTGGGGCA 5160
Db      5155 CAGCCCGACATCTACCCAAATTTTGAAGCTGTCAAGGGCTGCCACTTGTGGGGCA 5214
Qy      5161 GCGAATATTGAAGAGGGGCTCTGCACTGTGACGCCATGCCCGCAAAAGTGTCTATCTC 5220
Db      5215 GCGAATATTGAAGAGGGGCTCTGCACTGTGACGCCATGCCCGCAAAAGTGTCTATCTC 5274
Qy      5221 CGCTACAAAGAAACCTCAGCAATATCTGCATCCGAGAAAGATGAGACCTCAGAGCCC 5280
Db      5275 CGCTACAAAGAAACCTCAGCAATATCTGCATCCGAGAAAGATGAGACCTCAGAGCCC 5334
Qy      5281 TGCAGCTGTATCCACTTCAACCAATTAAGTATCTCATTTGGAACCAATAATTCTACGA 5340
Db      5335 TGCAGCTGTATCCACTTCAACCAATTAAGTATCTCATTTGGAACCAATAATTCTACGA 5394
Qy      5341 ATCGAATAGAAGAGTACAGCTGCGAGGAATTTCTGATTAAGATGACATTTCTTGGCA 5400
Db      5395 ATCGAATAGAAGAGTACAGCTGCGAGGAATTTCTGATTAAGATGACATTTCTTGGCA 5454
Qy      5401 CTGTGCTGTGTTGCGGCTCTTCAACAGCTTCCGTCTCAATCGAGAGGTGAACAGC 5460
Db      5455 CTGTGCTGTGTTGCGGCTCTTCAACAGCTTCCGTCTCAATCGAGAGGTGAACAGC 5514
Qy      5461 GCGAGGCGAGAGAGATCTTGTGCTGTGTTTCAAGAAATTTGAGTGTTCGTGATTC 5520
Db      5515 GCGAGGCGAGAGAGATCTTGTGCTGTGTTTCAAGAAATTTGAGTGTTCGTGATTC 5574
Qy      5521 TACGGAAGAGCTAGCGGACAGACATCTCAAGTGAAGTGTGCTTACCTTGGCTTGGC 5580
Db      5575 TACGGAAGAGCTAGCGGACAGACATCTCAAGTGAAGTGTGCTTACCTTGGCTTGGC 5634
Qy      5581 TACAGAGAACCTATCTGTGTTGACCACTTCAACTCACTGGAAGTAAATTGAGATCCAG 5640

```

```

Db      5635 TACAGAGAACCTATCTGTGTTGTGACCCACTTCAATCTCAAGTAATTGAGATCCAG 5694
Qy      5641 GCAAGCTCTCTCAGAGAGAACCTCTGCCGAGCGTACCTGACATCTCCGAACCCGCGTAC 5700
Db      5695 GCAAGCTCTCTCAGAGAGAACCTCTGCCGAGCGTACCTGACATCTCCGAACCCGCGTAC 5754
Qy      5701 CTGGGCGCTGCGCAATTTCTTCAAGAGCAATTTACTTGGCGTCTCTATACAGATTAATTA 5760
Db      5755 CTGGGCGCTGCGCAATTTCTTCAAGAGCAATTTACTTGGCGTCTCTATACAGATTAATTA 5814
Qy      5761 AGGTCATTTGTGCGCAAGAGAAACCTGTGAAGAGTCCGACCTGAACACACCGGAGC 5820
Db      5815 AGGTCATTTGTGCGCAAGAGAAACCTGTGAAGAGTCCGACCTGAACACACCGGAGC 5874
Qy      5821 CCGTCACTCTCCGACAGACCCCAACAGAGAGGCCCACTGTAACAGACATC 5880
Db      5875 CCGTCACTCTCCGACAGACCCCAACAGAGAGGCCCACTGTAACAGACATC 5934
Qy      5881 ACCAAGCGCTGGCTCTCCAGCCCAAGCGCGCCCGAAGGCCCAAGCCCGGAGAGCCA 5940
Db      5935 ACCAAGCGCTGGCTCTCCAGCCCAAGCGCGCCCGAAGGCCCAAGCCCGGAGAGCCA 5994
Qy      5941 AGCAGACCCCAAGCTACCGCGAGGGGCGGACCGAGCTGGCGAGGCAAGTCTCCTGGC 6000
Db      5995 AGCAGACCCCAAGCTACCGCGAGGGGCGGACCGAGCTGGCGAGGCAAGTCTCCTGGC 6054
Qy      6001 CGCCCCCTGGAGAGAGAAAGTCCCGCGGATCTACGACCGCGAGAGAGCGGTCC 6060
Db      6055 CGCCCCCTGGAGAGAGAAAGTCCCGCGGATCTACGACCGCGAGAGAGCGGTCC 6114
Qy      6061 CCGCGAGGCTGTTTGAAGACAGACAGAGGGGCGGCTGCTTCCGGAGCGCTGAGAGACC 6120
Db      6115 CCGCGAGGCTGTTTGAAGACAGACAGAGGGGCGGCTGCTTCCGGAGCGCTGAGAGACC 6174
Qy      6121 CCGCTGTCCAGGTGAACAAAGTCTGAGACAGTCTTCAGTATTA 6165
Db      6175 CCGCTGTCCAGGTGAACAAAGTCTGAGACAGTCTTCAGTATTA 6219

RESULT 4
AX671105      6165 bp      DNA      linear      PAT 27-MAR-2003
LOCUS      AX671105
DEFINITION      Sequence 1 from Patent WO03004523.
ACCESSION      AX671105
VERSION      AX671105.1 GI:29329569
KEYWORDS
SOURCE
ORGANISM      Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1.
Zhu, Z.
Regulation of human cIiron rho/rac-interacting kinase
Patent: WO 03004523-A 1 16--JAN-2003;
Bayer Aktiengesellschaft (DE)
FEATURES
source
1..6165
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match      99.9%; Score 6160.2; DB 6; Length 6165;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 6162; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
Qy      1 ATGTTGAAGTTCAATATGAGAGCGGAAATCTTTGAGTGTGTGCTGTAACCATTT 60
Db      1 ATGTTGAAGTTCAATATGAGAGCGGAAATCTTTGAGTGTGTGCTGTAACCATTT 60
Qy      61 GCCAGCGGGGCTCCAGGCGTGAATCTGTTCTTCAAGGGGAAACCAACCTTTATGACTCA 120
Db      61 GCCAGCGGGGCTCCAGGCGTGAATCTGTTCTTCAAGGGGAAACCAACCTTTATGACTCA 120

```


QY	121	CAGCAGATGTCCTCTTTTCCCGAGAAAGGATATTAGATGCCCTTGTGTTCTTTGAA	180
Db	121	CAGCAGATGTCCTCTCTTTCCCGAGAAAGGATATTAGATGCCCTTGTGTTCTTTGAA	180
QY	181	GAAATGCATCAGCCGCTCTGTATGAAGATTAGACCGTAGACCACTTGTCCGAAAGAT	240
Db	181	GAAATGCATCAGCCGCTCTGTATGAAGATTAGACCGTAGACCACTTGTCCGAAAGAT	240
QY	241	TCCGACACCAATAGCTGAGTTACAGGAGCTCCAGCCTTGGCGAAAAGSACTTCGAAATCAGA	300
Db	241	TCCGACACCAATAGCTGAGTTACAGGAGCTCCAGCCTTGGCGAAAAGSACTTCGAAATCAGA	300
QY	301	AGCTCTTAGGTTGTGTCACTTGTCTGTAAGTCAGGTGTGAAGAGAAAGCAACCGGG	360
Db	301	AGCTCTTAGGTTGTGTCACTTGTCTGTAAGTCAGGTGTGAAGAGAAAGCAACCGGG	360
QY	361	GACATCTATGCTATGAAGAAGTATGAAGAAGGCTTATTGGCCCAAGACAGATTTC	420
Db	361	GACATCTATGCTATGAAGAAGTATGAAGAAGGCTTATTGGCCCAAGACAGATTTC	420
QY	421	TTTTTTGAGAAAGCGGACATATATCTGGAAGCACAGCCGCTGATCCCCCAATTA	480
Db	421	TTTTTTGAGAAAGCGGACATATATCTGGAAGCACAGCCGCTGATCCCCCAATTA	480
QY	481	CAGTATGCCCTTTCAGAGCAAAAATCACCTTATCTGTATCATGAATATCAGCTGAGGG	540
Db	481	CAGTATGCCCTTTCAGAGCAAAAATCACCTTATCTGTATCATGAATATCAGCTGAGGG	540
QY	541	GACTTGCTGTCACTTTTGAATAGATATGAGAACCACTTAAATGTAACCTGATACAGTTT	600
Db	541	GACTTGCTGTCACTTTTGAATAGATATGAGAACCACTTAAATGTAACCTGATACAGTTT	600
QY	601	TACCTAGCTGAGCTGATTTTGGCTGTTACAGCCGTCATCTGATGGGATAGCTGATCAGA	660
Db	601	TACCTAGCTGAGCTGATTTTGGCTGTTACAGCCGTCATCTGATGGGATAGCTGATCAGA	660
QY	661	GACATCAAGCCCTGAGAACATTTCTGTGACCGCACAGSACATCAATCAAGCTGTGATTTT	720
Db	661	GACATCAAGCCCTGAGAACATTTCTGTGACCGCACAGSACATCAATCAAGCTGTGATTTT	720
QY	721	GGATCTGCGCGAAAATGAATTCAAAACAAGATGGGAATGCCAAATCTCCGATTTGGAGC	780
Db	721	GGATCTGCGCGAAAATGAATTCAAAACAAGATGGGAATGCCAAATCTCCGATTTGGAGC	780
QY	781	CCAATTTACATAGCTCTCTGAAGTGTGTACTGTATGAAACGGGAGATGAAAGGACCTTAC	840
Db	781	CCAATTTACATAGCTCTCTGAAGTGTGTACTGTATGAAACGGGAGATGAAAGGACCTTAC	840
QY	841	GAGCTGAGCTGTGACTGTGTGTCAGTGGCGGATGTCCTTATGAGATATTTATGGAGAA	900
Db	841	GAGCTGAGCTGTGACTGTGTGTCAGTGGCGGATGTCCTTATGAGATATTTATGGAGAA	900
QY	901	TCCCCCTTCGAGAGGAACTCTGCGCAAACTTCAATTAACATTATGAAATTTCCAGCGG	960
Db	901	TCCCCCTTCGAGAGGAACTCTGCGCAAACTTCAATTAACATTATGAAATTTCCAGCGG	960
QY	961	TTTTTGAATTTCCAGATGACCCCAAGATGACAGCTGACTTTCTTGAATCTGATTCAAAGC	1020
Db	961	TTTTTGAATTTCCAGATGACCCCAAGATGACAGCTGACTTTCTTGAATCTGATTCAAAGC	1020
QY	1021	TTGTTGTGCGGCGAAGAAAGAGAGACTGAAGTTTGAAGGCTTTGTGTGCATCTTTCTTC	1080
Db	1021	TTGTTGTGCGGCGAAGAAAGAGAGACTGAAGTTTGAAGGCTTTGTGTGCATCTTTCTTC	1080
QY	1081	TCTAAATTTGACTGGAACAACATTGTAACCTCTCTCCCTCGTTGTTCCACCTCTCAAG	1140
Db	1081	TCTAAATTTGACTGGAACAACATTGTAACCTCTCTCCCTCGTTGTTCCACCTCTCAAG	1140
QY	1141	TCTGACGATGACACTCCAAATTTTGTATGAACACAGAAAGAAATGTGGGTTTCACTCT	1200
Db	1141	TCTGACGATGACACTCCAAATTTTGTATGAACACAGAAAGAAATGTGGGTTTCACTCT	1200

QY	1201	CCGTCAGCTAGACCCCTTCAAGCTTCTCCGGTGAAGATCTGCCGTTTGTCGGGCTTTTCG	1250
Db	1201	CCGTGCCAGCTGAGACCCCTCAGGCTTCTCCGGGTGMAAATCGCGCTTGTCGGGCTTTTTCG	1250
QY	1261	TACAGCAAGGACCTGGGGATCTCTGGTATGATCTGAGCTGTGTGTGGGGCTTGAATCCC	1320
Db	1261	TACAGCAAGGACCTGGGGATCTTGTGTGATCTGATGCTGTGTGTGTGGGGCTTGAATCCC	1320
QY	1321	CCTGCGCAAGACTAGACTCCATGTGAAAAAGAACTTCTCATCAAAAGCAAAAGACTCAAGAC	1380
Db	1321	CCTGCGCAAGACTAGACTCCATGTGAAAAAGAACTTCTCATCAAAAGCAAAAGACTCAAGAC	1380
QY	1381	TCTCAGGCAAGTGTCAACAGATGAGACGAAATGACCCGGTTACATCCGAGAGTGTCA	1440
Db	1381	TCTCAGGCAAGTGTCAACAGATGAGACGAAATGACCCGGTTACATCCGAGAGTGTCA	1440
QY	1441	GAGGTGAGAGGTGTGCTTACATCAAGAAGGTGAGTGAAGGCTGTGAGACTCAAGA	1500
Db	1441	GAGGTGAGAGGTGTGCTTACATCAAGAAGGTGAGTGAAGGCTGTGAGACTCAAGA	1500
QY	1501	TCCCTCCTCGAGCAGACCTTGCTACCTACATCAAGAAATGCAATAGCACTTAAACGAAGT	1560
Db	1501	TCCCTCCTCGAGCAGACCTTGCTACCTACATCAAGAAATGCAATAGCACTTAAACGAAGT	1560
QY	1561	TTGGAGCAGACCGATGAGAGTGTCCGAGAGGATGACAAAGCAGCTTCATCTTCAT	1620
Db	1561	TTGGAGCAGACCGATGAGAGTGTCCGAGAGGATGACAAAGCAGCTTCATCTTCAT	1620
QY	1621	GATATCAGAGGACGAGCCGGAACTCTCAAGAAATCAAAAGCAGAGTACCGAGCTCAA	1680
Db	1621	GATATCAGAGGACGAGCCGGAACTCTCAAGAAATCAAAAGCAGAGTACCGAGCTCAA	1680
QY	1681	GTGGAAGAAATGAGAGTTGATGATGAATCAGTTGGAAGAGATCTTGTCTCAGCAAGAGA	1740
Db	1681	GTGGAAGAAATGAGAGTTGATGATGAATCAGTTGGAAGAGATCTTGTCTCAGCAAGAGA	1740
QY	1741	CGAGATGATCTCTACGAATCTGAGCTGAGAGAGTCTCGGCTTGCTGTGAAGATTCAG	1800
Db	1741	CGAGATGATCTCTACGAATCTGAGCTGAGAGAGTCTCGGCTTGCTGTGAAGATTCAG	1800
QY	1801	CGGAAGCGACAGAAATGTCAGCATAAATCTGTAAGGCTTAAGATCAAGGAAAGCTTGA	1860
Db	1801	CGGAAGCGACAGAAATGTCAGCATAAATCTGTAAGGCTTAAGATCAAGGAAAGCTTGA	1860
QY	1861	GTGGGAGATATGCGAAATCTGGAAGATGCAATCTGATGAGCAGCAGCTCAAAATTCAGAG	1920
Db	1861	GTGGGAGATATGCGAAATCTGGAAGATGCAATCTGATGAGCAGCAGCTCAAAATTCAGAG	1920
QY	1921	CTCCAAAGAAATCTGAGAAAGGCTGTAAAGCCAGCAGCGAGGCGCACCGAGCTGCTCAG	1980
Db	1921	CTCCAAAGAAATCTGAGAAAGGCTGTAAAGCCAGCAGCGAGGCGCACCGAGCTGCTCAG	1980
QY	1981	AATATCCGCCAGGCAAAAGAGCGAGCCGAGAGGAGCTGAGAACTGTCAAGCCGAGAG	2040
Db	1981	AATATCCGCCAGGCAAAAGAGCGAGCCGAGAGGAGCTGAGAACTGTCAAGCCGAGAG	2040
QY	2041	GATTCCTCTGAAGCATACGAAABAAGCTGTGAACTGTGAAGAACCGCGCATTTCTGTG	2100
Db	2041	GATTCCTCTGAAGCATACGAAABAAGCTGTGAACTGTGAAGAACCGCGCATTTCTGTG	2100
QY	2101	GAGAAACAAGTAAAGAGACTAGAGACCATGAGCGTATGAGAAAAACAATTAAGATATAC	2160
Db	2101	GAGAAACAAGTAAAGAGACTAGAGACCATGAGCGTATGAGAAAAACAATTAAGATATAC	2160
QY	2161	ATCCAGACAAATATCCCAACAGATCCAGACATGAGCTGATAAATTTCTGAGCTTGAAGAG	2220
Db	2161	ATCCAGACAAATATCCCAACAGATCCAGACATGAGCTGATAAATTTCTGAGCTTGAAGAG	2220
QY	2221	AAACATCGGGAGGCGCCAAAGTCTCAGCCGACAGACTAGAAAGGCACTGAAACAAGAAAG	2280
Db	2221	AAACATCGGGAGGCGCCAAAGTCTCAGCCGACAGACTAGAAAGGCACTGAAACAAGAAAG	2280
QY	2281	CAGCATATATGAGAAAAAGATTAAAGTGTGAGCAATCAGATTAAGAAAGACTTGGCTGAC	2340

```
Db 2281 CAGACCTATGAGAAAAGATTAAAGTTTGGACATATCAATTAAGAAAGACCTGAGCTGAC 2340
Qy 2341 AAGAGACACTGAGGAA CATGATGCAAGACACGAGGAGAGGCCATGAGAGGGGAAA 2400
Db 2341 AAGGAGACACTGAGGAA CATGATGCAAGACACGAGGAGAGGCCATGAGAGGGGAAA 2400
Qy 2401 ATTCTGCGGAAACGAGAGGGCATATCAATGCTATGATTTCCAGATTCAGATTCAGATTCCTGGAA 2460
Db 2401 ATTCTGCGGAAACGAGAGGGCATATCAATGCTATGATTTCCAGATTCAGATTCCTGGAA 2460
Qy 2461 CAGAGATTGTGGAATGCTGTGAAGCCAAATTAACCTTGACGAAATAGCAAGCTTTTAAAC 2520
Db 2461 CAGAGATTGTGGAATGCTGTGAAGCCAAATTAACCTTGACGAAATAGCAAGCTTTTAAAC 2520
Qy 2521 CAAAGGAACTGAAAGGCCCAAGAAAGATGATTTCTGAACCTCAGGCAACAGAAATTTTAC 2580
Db 2521 CAAAGGAACTGAAAGGCCCAAGAAAGATGATTTCTGAACCTCAGGCAACAGAAATTTTAC 2580
Qy 2581 CTGAGACACAGGCTGGGAAAGTTGAGAGGCCCAAGAACTGAGAGAGCAGCTGGAG 2640
Db 2581 CTGAGACACAGGCTGGGAAAGTTGAGAGGCCCAAGAACTGAGAGAGCAGCTGGAG 2640
Qy 2641 AAGATCAGCCACCAAGACCAACAGTGAACAAGATCGGCTGGAACTGAGACAAAGATTG 2700
Db 2641 AAGATCAGCCACCAAGACCAACAGTGAACAAGATCGGCTGGAACTGAGACAAAGATTG 2700
Qy 2701 CGGAGGTCAGTCTAGACACAGAGACAGAGAACTGAGCTCAAGCCAGCTCACAGAG 2760
Db 2701 CGGAGGTCAGTCTAGAGACACAGAGACAGAGAACTGAGAGCTCAAGCCAGCTCACAGAG 2760
Qy 2761 CTACAGCTCTCCCTGAGAGAGCGCGAGTCAAGTTGACAGCCCTGACGCGTGCACGCGCG 2820
Db 2761 CTACAGCTCTCCCTGAGAGAGCGCGAGTCAAGTTGACAGCCCTGACGCGTGCACGCGCG 2820
Qy 2821 GCCCTGAGAGGCGAGCTTGCGCGAGGGAAGACAGAGTGGAAAGACCAAGAGAAAGCT 2880
Db 2821 GCCCTGAGAGGCGAGCTTGCGCGAGGGAAGACAGAGTGGAAAGACCAAGAGAAAGCT 2880
Qy 2881 GAAAGAGAGATCCAGGCACTCACGAGCAATAGAGATGAATAATCCAGGCGAAATTTGATGCT 2940
Db 2881 GAAAGAGAGATCCAGGCACTCACGAGCAATAGAGATGAATAATCCAGGCGAAATTTGATGCT 2940
Qy 2941 CTTCTGTAACAGCTGTATCTGTATATCAAGACTTGAGAGAGCAAGCTTAACCAAGCTGAC 3000
Db 2941 CTTCTGTAACAGCTGTATCTGTATATCAAGACTTGAGAGAGCAAGCTTAACCAAGCTGAC 3000
Qy 3001 GACAAAGCTGAATCTCAACCAACCAAACTTCTATGTCAAACAACTCGATGAGGCTTCT 3060
Db 3001 GACAAAGCTGAATCTCAACCAACCAAACTTCTATGTCAAACAACTCGATGAGGCTTCT 3060
Qy 3061 GCGCGCCACGAGAGATTGTACACTGCGAAGTGAAGTGAACCATCTCCGCGCGGAGATC 3120
Db 3061 GCGCGCCACGAGAGATTGTACACTGCGAAGTGAAGTGAACCATCTCCGCGCGGAGATC 3120
Qy 3121 ACGGAAACGAGATGAGCTTACACGCGAAGAACAAACGATGAGGCTCTGAAGACCAAG 3180
Db 3121 ACGGAAACGAGATGAGCTTACACGCGAAGAACAAACGATGAGGCTCTGAAGACCAAG 3180
Qy 3181 TGCACCATGCTGAGAGAAACAGATCAATGATTTTGAGAGCCCTTAACCATGAGCTGTAGAA 3240
Db 3181 TGCACCATGCTGAGAGAAACAGATCAATGATTTTGAGAGCCCTTAACCATGAGCTGTAGAA 3240
Qy 3241 AAAGAGCGCAGTGGAGGCTTGAGAGAGCTCTTGAGTATGAGAAATCCCAAGTTTGAAG 3300
Db 3241 AAAGAGCGCAGTGGAGGCTTGAGAGAGCTCTTGAGTATGAGAAATCCCAAGTTTGAAG 3300
Qy 3301 TGTGCGGTTTGAAGCTGAGAGATGCTTGAACACCGAGAAAACAGAGCAGGCGAGAGCC 3360
Db 3301 TGTGCGGTTTGAAGCTGAGAGATGCTTGAACACCGAGAAAACAGAGCAGGCGAGAGCC 3360
Qy 3361 GATCAGCGGATACCGAGTCTTCCAGGCTGTGAGAGCTGGCAGTGAAGAGCAAGAGCT 3420
Db 3361 GATCAGCGGATACCGAGTCTTCCAGGCTGTGAGAGCTGGCAGTGAAGAGCAAGAGCT 3420
Qy 3421 GAGATTCTCGCTGTGACAGAGCTCTTCAAAAGAGCAAGAGCTGAAGGCCAGAGCTCTCT 3480
Db 3421 GAGATTCTCGCTGTGACAGAGCTCTTCAAAAGAGCAAGAGCTGAAGGCCAGAGCTCTCT 3480
Qy 3481 GACAGCTCAATGACTTGAAGAGAGAGATGCTATGCTTGAATGAATGAATGCCGGAAGCTTA 3540
Db 3481 GACAGCTCAATGACTTGAAGAGAGAGATGCTATGCTTGAATGAATGAATGCCGGAAGCTTA 3540
Qy 3541 CAGAGAGCTGAGACTGAACGAGAGCTGAACAGAGGCTTCTGAGAGAGCAAGCCAAA 3600
Db 3541 CAGAGAGCTGAGACTGAACGAGAGCTGAACAGAGGCTTCTGAGAGAGCAAGCCAAA 3600
Qy 3601 TTAACAGAGCAGATGAGACTTGCAAGAAAATACATTTTCCGTCTGACTCAAGACTGCA 3660
Db 3601 TTAACAGAGCAGATGAGACTTGCAAGAAAATACATTTTCCGTCTGACTCAAGACTGCA 3660
Qy 3661 GAACTCTAGATGGGCTGATCTATGAGACAGAAAAGAGATGATCTTGAGATCAAGCTG 3720
Db 3661 GAACTCTAGATGGGCTGATCTATGAGACAGAAAAGAGATGATCTTGAGATCAAGCTG 3720
Qy 3721 GAAACATTCAGGTTCTTAATCTCAATGAAGAGTGAAGTGAAGGCACTATTTCTCA 3780
Db 3721 GAAACATTCAGGTTCTTAATCTCAATGAAGAGTGAAGTGAAGGCACTATTTCTCA 3780
Qy 3781 CAAACCAAATCTCATTTGATTTTGTCAAGCCAAATGSAACCACTCTGTAATAAGAAAAG 3840
Db 3781 CAAACCAAATCTCATTTGATTTTGTCAAGCCAAATGSAACCACTCTGTAATAAGAAAAG 3840
Qy 3841 GTTCTCTGCAATCAATGAGCTGAAGCTGGCCCTGAGAGAGAGAAAGCTGCTGTGCA 3900
Db 3841 GTTCTCTGCAATCAATGAGCTGAAGCTGGCCCTGAGAGAGAGAAAGCTGCTGTGCA 3900
Qy 3901 GAGCTTGAAGAGAGCCCTTTGAGAAAGCCGCAATGAGCTCCGCTCCGCGGAGAGAGCT 3960
Db 3901 GAGCTTGAAGAGAGCCCTTTGAGAAAGCCGCAATGAGCTCCGCTCCGCGGAGAGAGCT 3960
Qy 3961 GCCACCGCAAGAACAGAGACCAACCAACCAATCCAGCCAGCCAGCCAGCGAGAGCAG 4020
Db 3961 GCCACCGCAAGAACAGAGACCAACCAACCAATCCAGCCAGCCAGCGAGAGCAG 4020
Qy 4021 ATGCGCATGCTCGCATGCTGCGGTGCGCAGAGACACAGCCAGTGCATGAGCCTGTG 4080
Db 4021 ATGCGCATGCTCGCATGCTGCGGTGCGCAGAGACACAGCCAGTGCATGAGCCTGTG 4080
Qy 4081 GCCCGCCATCCAGCGCGAGAAAGAGTCTTAATCCAGAGAAATTTAGTGGGCTCTT 4140
Db 4081 GCCCGCCATCCAGCGCGAGAAAGAGTCTTAATCCAGAGAAATTTAGTGGGCTCTT 4140
Qy 4141 AAGGAAACGATCACCAATATTTCTCAACCGATTCAACGTAGACTGAACATGCGAGCC 4200
Db 4141 AAGGAAACGATCACCAATATTTCTCAACCGATTCAACGTAGACTGAACATGCGAGCC 4200
Qy 4201 ACAAAGTGTGCTGTGTGTCTGATACCGTGACCTTTGAGCGCAAGCAATCCAAATGTCTC 4260
Db 4201 ACAAAGTGTGCTGTGTGTCTGATACCGTGACCTTTGAGCGCAAGCAATCCAAATGTCTC 4260
Qy 4261 GAATGTCAAGTATGTGTCAACCCCAAGTGTCTCAAGTGTCTTCCAGCCACCTGCGGCTTG 4320
Db 4261 GAATGTCAAGTATGTGTCAACCCCAAGTGTCTCAAGTGTCTTCCAGCCACCTGCGGCTTG 4320
Qy 4321 CCTGCTGAATATGCAACAATCTTCAACGAGGCTTCTGCGCGTGAACAAATGAACCTCCCA 4380
Db 4321 CCTGCTGAATATGCAACAATCTTCAACGAGGCTTCTGCGCGTGAACAAATGAACCTCCCA 4380
Qy 4381 GGTCTCAAGACCAAGAGCCCAAGCAGCTTGAACCTTGAAGGCTGATGAAGGTGCC 4440
Db 4381 GGTCTCAAGACCAAGAGCCCAAGCAGCTTGAACCTTGAAGGCTGATGAAGGTGCC 4440
Qy 4441 AGGAATTAACAAACGAGACAGAGAGCTGGGACAGAAATACATTTGCTTGAAGGATCA 4500
Db 4441 AGGAATTAACAAACGAGACAGAGAGCTGGGACAGAAATACATTTGCTTGAAGGATCA 4500
```

QY 4501 AAAGTCTCTATTTATGCAATGAAAGCCAGAAAGCTGGACAGAGCCGGTGAAGATTT 4560
 DB 4501 AAAGTCTCTATTTATGCAATGAAAGCCAGAAAGCTGGACAGAGCCGGTGAAGATTT 4560
 QY 4561 GAGCTGTCCTTCCCGAGGGGATGTATTTCAATGTCGTCCTTGGTGTCTTCCGAATC 4620
 DB 4561 GAGCTGTCCTTCCCGAGGGGATGTATTTCAATGTCGTCCTTGGTGTCTTCCGAATC 4620
 QY 4621 GCAATATGAGCCAAAGAGATGTCTCCATACATCTGAAATGGAATCTCACCCGACACC 4680
 DB 4621 GCAATATGAGCCAAAGAGATGTCTCCATACATCTGAAATGGAATCTCACCCGACACC 4680
 QY 4681 ACCTGTGCGCCGGAGAAACCTCTTACTTCTGCTCCAGCTTCCCTGCAAAACAGCGC 4740
 DB 4681 ACCTGTGCGCCGGAGAAACCTCTTACTTCTGCTCCAGCTTCCCTGCAAAACAGCGC 4740
 QY 4741 TGGGTCAAGCCCTTAAATGATGTTGTGCAAGTGGAGAGATTTCTAGGAAAAAGCAGAA 4800
 DB 4741 TGGGTCAAGCCCTTAAATGATGTTGTGCAAGTGGAGAGATTTCTAGGAAAAAGCAGAA 4800
 QY 4801 GCTGATGCTAAATGCTTGAAGAACTCCCTGCTGAACTGAAAGGTGATGACCGTCTAGAC 4860
 DB 4801 GCTGATGCTAAATGCTTGAAGAACTCCCTGCTGAACTGAAAGGTGATGACCGTCTAGAC 4860
 QY 4861 ATGAATGCAAGCTGCGCTTCACTGACAGAGTGGTGTGGTGGCAACGAGAAAGGCTC 4920
 DB 4861 ATGAATGCAAGCTGCGCTTCACTGACAGAGTGGTGTGGTGGCAACGAGAAAGGCTC 4920
 QY 4921 TACGCTCTGATGTCTTGAAGAACTCCCTTAAACCATATCTCCAGGAATTGAAGCATCTTC 4980
 DB 4921 TACGCTCTGATGTCTTGAAGAACTCCCTTAAACCATATCTCCAGGAATTGAAGCATCTTC 4980
 QY 4981 CAAATTTATTTATGCAAGGACTGAGAAAGCTACTCATGATAGAGAGAAAGCGGGCA 5040
 DB 4981 CAAATTTATTTATGCAAGGACTGAGAAAGCTACTCATGATAGAGAGAAAGCGGGCA 5040
 QY 5041 CTGTGCTTGTGACGCTGAGAAAGTGAACAGTCCCTGCGCCAGTCCACTGCTGCGC 5100
 DB 5041 CTGTGCTTGTGACGCTGAGAAAGTGAACAGTCCCTGCGCCAGTCCACTGCTGCGC 5100
 QY 5101 CAGCCGACATCTCAACCAATTTTGAAGCTGTCAAGGGCTGCACTTGTGTGGGGCA 5160
 DB 5101 CAGCCGACATCTCAACCAATTTTGAAGCTGTCAAGGGCTGCACTTGTGTGGGGCA 5160
 QY 5161 GGGAAAGTTAGAAAGGGGCTGTGATCTGTGACGCAATGCCGCAAGATGCTCATCTC 5220
 DB 5161 GGGAAAGTTAGAAAGGGGCTGTGATCTGTGACGCAATGCCGCAAGATGCTCATCTC 5220
 QY 5221 CGCTACACGAAACCTCAGCAAAATCTGCATCCGGAAGAGATAGAGACTCAGAGCCC 5280
 DB 5221 CGCTACACGAAACCTCAGCAAAATCTGCATCCGGAAGAGATAGAGACTCAGAGCCC 5280
 QY 5281 TGCAGCTGATCCATCTTCAACCAATTTAGATCTCTTCAATCGAGCAATTAATCTAGAA 5340
 DB 5281 TGCAGCTGATCCATCTTCAACCAATTTAGATCTCTTCAATCGAGCAATTAATCTAGAA 5340
 QY 5341 ATGACATGAAAGATGACGCTGAGAGATCTCTGATTAAGATGACATCTCTTGGCA 5400
 DB 5341 ATGACATGAAAGATGACGCTGAGAGATCTCTGATTAAGATGACATCTCTTGGCA 5400
 QY 5401 CCGTGTGTTTGGCGGCTCTTCCAAACAGCTTCCCTGCTCAATCGAGCAAGTAAAGC 5460
 DB 5401 CCGTGTGTTTGGCGGCTCTTCCAAACAGCTTCCCTGCTCAATCGAGCAAGTAAAGC 5460
 QY 5461 GCAAGGAGAGAGAGATCTTGTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 5520
 DB 5461 GCAAGGAGAGAGAGATCTTGTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 5520
 QY 5521 TACGGAAGAGTACGCGCAGAGATCTCAAGTGAAGTGTGCTTACCTTTGGCTTTGGC 5580
 DB 5521 TACGGAAGAGTACGCGCAGAGATCTCAAGTGAAGTGTGCTTACCTTTGGCTTTGGC 5580

QY 5581 TACAGGAACCTATCTGTTTGTGACCCACTTAACTCACTCGAAGTAAATTGATCCAG 5640
 DB 5581 TACAGGAACCTATCTGTTTGTGACCCACTTAACTCACTCGAAGTAAATTGATCCAG 5640
 QY 5641 GCAAGCTCCACAGAGAACCCCTGCGCAGAGCTA CTTGGAATCCCGAAGCCGCTAC 5700
 DB 5641 GCAAGCTCCACAGAGAACCCCTGCGCAGAGCTA CTTGGAATCCCGAAGCCGCTAC 5700
 QY 5701 CTGGGCTTGCATTTCTCAGAGAGATTTACTTGGCGTCTCATACAGATTAATTA 5760
 DB 5701 CTGGGCTTGCATTTCTCAGAGAGATTTACTTGGCGTCTCATACAGATTAATTA 5760
 QY 5761 AGGCTATTTGCTGCAAGGAAACCTCTGTAAGAGATCGGCACTGAACACACCGGGGC 5820
 DB 5761 AGGCTATTTGCTGCAAGGAAACCTCTGTAAGAGATCGGCACTGAACACACCGGGGC 5820
 QY 5821 CCGTCAACCTCCGAGAGAGCCCAACAGAGAGCCCAACCGAAGTCAACAGACATC 5880
 DB 5821 CCGTCAACCTCCGAGAGAGCCCAACAGAGAGCCCAACCGAAGTCAACAGACATC 5880
 QY 5881 ACCAAGCGCTGCTCTCAGCCAGCGCCCGCAAGGCCCAAGCCACCGCAGAGCCA 5940
 DB 5881 ACCAAGCGCTGCTCTCAGCCAGCGCCCGCAAGGCCCAAGCCACCGCAGAGCCA 5940
 QY 5941 AGCACAACCCACCGCTACCGAGAGGGGCGGACCGAGCTGCGAGGACAAATGTTCTTGGC 6000
 DB 5941 AGCACAACCCACCGCTACCGAGAGGGGCGGACCGAGCTGCGAGGACAAATGTTCTTGGC 6000
 QY 6001 CGGCTTGTGAGAGAGAGAGAGTCCCCGCGGATCTCAAGCAAGCGAGAGAGCGGTTC 6060
 DB 6001 CGGCTTGTGAGAGAGAGAGAGTCCCCGCGGATCTCAAGCAAGCGAGAGAGCGGTTC 6060
 QY 6061 CCCGAGAGCTGTTTGAAGACAGACAGAGGGGCGGCTGCTGCGGAGCGGTGAGAGCC 6120
 DB 6061 CCCGAGAGCTGTTTGAAGACAGACAGAGGGGCGGCTGCTGCGGAGCGGTGAGAGCC 6120
 QY 6121 CCGCTGTCCAGGTGAACAAAGTCTGGAGCAAGTCTTCAATATA 6165
 DB 6121 CCGCTGTCCAGGTGAACAAAGTCTGGAGCAAGTCTTCAATATA 6165

RESULT 5
 AX671108
 LOCUS AX671108 8603 bp DNA linear PAT 27-MAR-2003
 DEFINITION Sequence 4 from Patent WO0304523.
 ACCESSION AX671108
 VERSION AX671108.1 GI:29329570
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE
 AUTHORS Zhu Z.
 TITLE Regulation of human cItron rho/rac-interacting kinase
 JOURNAL Patent: WO 0304523-A 4 16-JAN-2003;
 Bayer Aktiengesellschaft (DE)
 FEATURES
 source
 1..8603
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"
 ORIGIN
 Query Match 99.9%; Score 6160.2; DB 6; Length 8603;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 6162; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 1 ATGTTGAAGTTCAATATGAGAGCGGGAATCTTTGGATCTGATGCTGTAACCATTT 60
 DB 1 ATGTTGAAGTTCAATATGAGAGCGGGAATCTTTGGATCTGATGCTGTAACCATTT 60
 QY 61 GCCAGCGGCGCTCCAGGCTGAATCTGTTCTTCAAGGGAACCACTTTATGACTCAA 120

Db	61	GCAGCGCGGCCCTCCAGGCTGGAATCTGTTCTTCAGGGGAAACACCCCTTATGACTCAA	120
QY	121	CAGCAGATGTCCTCTCTTTCCGAGAAAGGATATTTAGATGCCCTCTTGTCTCTTTGAA	180
Db	121	CAGCAGATGTCCTCTCTTTCCGAGAAAGGATATTTAGATGCCCTCTTGTCTCTTTGAA	180
QY	181	GAAATGAGTCAAGCTGCTCTGATGAAAGTAAACAGTGAAGCACTTTGTCCGAAAGTAT	240
Db	181	GAAATGAGTCAAGCTGCTCTCTGATGAAAGTAAACAGTGAAGCACTTTGTCCGAAAGTAT	240
QY	241	TCGACACCAATAGCTAGTTTACAGAGCTCCAGCCTTCGCAAAAGCACTTCAGAGTCAAG	300
Db	241	TCGACACCAATAGCTAGTTTACAGAGCTCCAGCCTTCGCAAAAGCACTTCAGAGTCAAG	300
QY	301	AGTCTTGATGTTGTGTCACTTTGCTGAAGTGCAGGTGGTAAAGAGAAAGCAACCGGG	360
Db	301	AGTCTTGATGTTGTGTGTCACTTTGCTGAAGTGCAGGTGGTAAAGAGAAAGCAACCGGG	360
QY	361	GACATCTATGCTATGAAAGTATGAAAGAAAGGCTTTATTTGCCCAAGAGCAGTTTCA	420
Db	361	GACATCTATGCTATGAAAGTATGAAAGAAAGGCTTTATTTGCCCAAGAGCAGTTTCA	420
QY	421	TTTTTTGAGAAAGAGCGGAAACATATATCTCGAAGCAAGCCCGTGAATCCGCCCAATTA	480
Db	421	TTTTTTGAGAAAGAGCGGAAACATATATCTCGAAGCAAGCCCGTGAATCCGCCCAATTA	480
QY	481	CAGTATGCTTTTCAAGACAAAATCACTTTATCTGGTCATGGAATATCAGCCTGAGGG	540
Db	481	CAGTATGCTTTTCAAGACAAAATCACTTTATCTGGTCATGGAATATCAGCCTGAGGG	540
QY	541	GACTTGCTGCTCATTTTGAATATGATATGAGGACCACTTATGATGAAACCTGATACATTT	600
Db	541	GACTTGCTGCTCATTTTGAATATGATATGAGGACCACTTATGATGAAACCTGATACATTT	600
QY	601	TACCTAGCTGACCTGATTTTGGCTGTTTCAACAGCTTCACTGATGGGATACGTGATGTA	660
Db	601	TACCTAGCTGACCTGATTTTGGCTGTTTCAACAGCTTCACTGATGGGATACGTGATGTA	660
QY	661	GACATCAAGCCTGAGAACATTTCTGTTGACCGCACAGGACACATCAAGCTGTGTGATTTT	720
Db	661	GACATCAAGCCTGAGAACATTTCTGTTGACCGCACAGGACACATCAAGCTGTGTGATTTT	720
QY	721	GATCTGCGCGCCAAAATGATTAACAAGATGGTGAATGCCAACTCCCGATTGGGACC	780
Db	721	GATCTGCGCGCCAAAATGATTAACAAGATGGTGAATGCCAACTCCCGATTGGGACC	780
QY	781	CCAGATTTCATGGGCTCTGAAGTGTGACTGTGATGAAACGGGATATGAAAGGCACTTAC	840
Db	781	CCAGATTTCATGGGCTCTGAAGTGTGACTGTGATGAAACGGGATATGAAAGGCACTTAC	840
QY	841	GAGCTGGACTGTAAGTGTGTCACTGAGGCGCTGATTTGCTATGATGATTTTATGGGAGA	900
Db	841	GAGCTGGACTGTAAGTGTGTCACTGAGGCGCTGATTTGCTATGATGATTTTATGGGAGA	900
QY	901	TCCCCCTTTCGAGAGGGAACCTCTGCCAAGACCTTCAATTAACATTATGAAATTTCCAGGG	960
Db	901	TCCCCCTTTCGAGAGGGAACCTCTGCCAAGACCTTCAATTAACATTATGAAATTTCCAGGG	960
QY	961	TTTTTGAATTTTCAGATGACCCCAAGATGACATGACTTTCTTGATCTGATTTCAAGC	1020
Db	961	TTTTTGAATTTTCAGATGACCCCAAGATGACATGACTTTCTTGATCTGATTTCAAGC	1020
QY	1021	TTGTTGTGCGGCCAGAAAGAGACTGAAGTTTGAAGGCTTTGTGTCATCTTTCTTC	1080
Db	1021	TTGTTGTGCGGCCAGAAAGAGACTGAAGTTTGAAGGCTTTGTGTCATCTTTCTTC	1080
QY	1081	TCATAAATTTGACTGGAACAACATTTCTGTAATCTCTCTCCCTCTGTTCCACCTCAAG	1140
Db	1081	TCATAAATTTGACTGGAACAACATTTCTGTAATCTCTCTCCCTCTGTTCCACCTCAAG	1140
QY	1141	TCATGACATGACCTTCCAATTTTATGTAACCAAGAAAGAAATTCGTGGGTTTCAATCTCT	1200
Db	1141	TCATGACATGACCTTCCAATTTTATGTAACCAAGAAAGAAATTCGTGGGTTTCAATCTCT	1200

Db	1141	TCGACGATGACACCTTCCTCAATTTTGAATGAACAGAGAAAGAAATTCGTGGGTTTCAATCCCTCT	1200
QY	1201	CCGTGCAAGCTGAGCCCTCTCAGGCTTCTCGGATGAAGAACTGCCGTTTGTGGGGTTTTCG	1260
Db	1201	CCGTGCAAGCTGAGCCCTCTCAGGCTTCTCGGATGAAGAACTGCCGTTTGTGGGGTTTTCG	1260
QY	1261	TAACGCAAGGCACTGGGGATTTCTTTGGTAGATCTGAAGTCTGTGTGTGGGTTCTGCACTCC	1320
Db	1261	TAACGCAAGGCACTGGGGATTTCTTTGGTAGATCTGAAGTCTGTGTGTGGGTTCTGCACTCC	1320
QY	1321	CCGCGCAAGACTGACTGCTCAATGAAAAAGAACTTCTCATCAAAAAGCAAAAGCTCAAGAC	1380
Db	1321	CCGCGCAAGACTGACTGCTCAATGAAAAAGAACTTCTCATCAAAAAGCAAAAGCTCAAGAC	1380
QY	1381	TCTCAGAGCAAGGTGTCACAAGATGAGCAGAAATGAACCGGTTTCAATCGAGAGTGTCA	1440
Db	1381	TCTCAGAGCAAGGTGTCACAAGATGAGCAGAAATGAACCGGTTTCAATCGAGAGTGTCA	1440
QY	1441	GAGGTGAGGCTGTGCTTGTGTCAAGAGAGGTGAGCTGAAGCCTTGAAGTCTCAGAGA	1500
Db	1441	GAGGTGAGGCTGTGCTTGTGTCAAGAGAGGTGAGCTGAAGCCTTGAAGTCTCAGAGA	1500
QY	1501	TCCCTCTCTGAGAGGAGACCTTGTCTACTCTCATCAAGAAATGCAAGTACTTAAAGCGAAGT	1560
Db	1501	TCCCTCTCTGAGAGGAGACCTTGTCTACTCTCATCAAGAAATGCAAGTACTTAAAGCGAAGT	1560
QY	1561	TTGAGCAAGCAACGAGTGAAGGTGTCCAGAGAGATGACAAAGACACTGCAGTCTTCCAT	1620
Db	1561	TTGAGCAAGCAACGAGTGAAGGTGTCCAGAGAGATGACAAAGACACTGCAGTCTTCCAT	1620
QY	1621	GATATCAGAGAGCAGAGCCGGAAGCTTCCAAAGAAATCAAAGAGCAGAGTACACAGGCTCAA	1680
Db	1621	GATATCAGAGAGCAGAGCCGGAAGCTTCCAAAGAAATCAAAGAGCAGAGTACACAGGCTCAA	1680
QY	1681	GTGGAAGAAATGAGGTGTAATGAATCAAGTTGGAAGAGAGATCTTGTCTCAGAGAGAGA	1740
Db	1681	GTGGAAGAAATGAGGTGTAATGAATCAAGTTGGAAGAGAGATCTTGTCTCAGAGAGAGA	1740
QY	1741	CGAGTGAATCTTACGAATCTGAGCTGAGAGAGTCTCGGCTTGTCTCTGAAGAAATTCAG	1800
Db	1741	CGAGTGAATCTTACGAATCTGAGCTGAGAGAGTCTCGGCTTGTCTCTGAAGAAATTCAG	1800
QY	1801	CGGAAAGCGACAGAAATGTCAGCATAACTGTTGAAGGCTTGAAGATCAAGGAAAGCCTGAA	1860
Db	1801	CGGAAAGCGACAGAAATGTCAGCATAACTGTTGAAGGCTTGAAGATCAAGGAAAGCCTGAA	1860
QY	1861	GTGGAGGAATATGCGAAACTGAGAGAAATCAATGCTGACACAGAGCTCAAAAATTCAGAG	1920
Db	1861	GTGGAGGAATATGCGAAACTGAGAGAAATCAATGCTGACACAGAGCTCAAAAATTCAGAG	1920
QY	1921	CTCCAAAGAGAAACTGAGAGAGGCTGTAAAAAGCCAGACCGAGAGCCACCGAGCTGCTCAG	1980
Db	1921	CTCCAAAGAGAAACTGAGAGAGGCTGTAAAAAGCCAGACCGAGAGCCACCGAGCTGCTCAG	1980
QY	1981	AATATCCGCCAGGCAAGAGCGAGCCGAGAGAGGAGCTGAGAAAGCTTCAGAACCGAGAG	2040
Db	1981	AATATCCGCCAGGCAAGAGCGAGCCGAGAGAGGAGCTGAGAAAGCTTCAGAACCGAGAG	2040
QY	2041	GATTCTTTCTGAAGGCACTCAAGAAAGAGCTGTGTGAAGCTGAGAGAAAGCCGCCATTTCTCG	2100
Db	2041	GATTCTTTCTGAAGGCACTCAAGAAAGAGCTGTGTGAAGCTGAGAGAAAGCCGCCATTTCTCG	2100
QY	2101	GAGAAACAAGGTAAAGAGACTGAGACCAATGAGCGTAGAGAAAACAGACTGAAGAGTGAAC	2160
Db	2101	GAGAAACAAGGTAAAGAGACTGAGACCAATGAGCGTAGAGAAAACAGACTGAAGAGTGAAC	2160
QY	2161	ATCCAGACAAATCCCAACAGATCCAGACAGATGAGCTGTATTAATTTCTGAGGCTCGAAGAG	2220
Db	2161	ATCCAGACAAATCCCAACAGATCCAGACAGATGAGCTGTATTAATTTCTGAGGCTCGAAGAG	2220
QY	2221	AAACATCGGAGGCGCCAAAGTCTCAGCCCAAGCACTTGAAGTGCACCTTGAACAGAAAGAG	2280
Db	2221	AAACATCGGAGGCGCCAAAGTCTCAGCCCAAGCACTTGAAGTGCACCTTGAACAGAAAGAG	2280

QY 2281 CAGCATATAGAGAAAAGATTAAAGTGTGGACAAATCAGATTAAGAAAAGACTGGCTGAC 2340
| | | | |
Db 2281 CAGCACTATAGAGAAAAGATTAAAGTGTGGACAAATCAGATTAAGAAAAGACTGGCTGAC 2340
QY 2341 AAGGAGCAGTGGAGCAATGATGCGAGACACGAGAGAGAGGCGCCATGGAAGGGGAAA 2400
| | | | |
Db 2341 AAGGAGCAGTGGAGCAATGATGCGAGACACGAGAGAGAGGCGCCATGGAAGGGGAAA 2400
QY 2401 ATTCTCAGCGAAACGAGAGCGATGATCAATGCTATGATTCCAGATCAAGATCCCTGGAA 2460
| | | | |
Db 2401 ATTCTCAGCGAAACGAGAGCGATGATCAATGCTATGATTCCAGATCAAGATCCCTGGAA 2460
QY 2461 CAGAGATTGTGAATCTGTCTGAAGCCAAATTAACCTTCAGCAAAATAGCAAGTCTTTTACC 2520
| | | | |
Db 2461 CAGAGATTGTGAATCTGTCTGAAGCCAAATTAACCTTCAGCAAAATAGCAAGTCTTTTACC 2520
QY 2521 CAAAAGAACTGAAAGGCGCAAGAAAGATGATTTTCTGAACCTCAGCAAAATAGCAAGTCTTTTACC 2580
| | | | |
Db 2521 CAAAAGAACTGAAAGGCGCAAGAAAGATGATTTTCTGAACCTCAGCAAAATAGCAAGTCTTTTACC 2580
QY 2581 CTGGAGACACAGGCTGGAGAGTTGGAGGCCAGAAACGAAACTGGAGAGCAGCTGGAG 2640
| | | | |
Db 2581 CTGGAGACACAGGCTGGAGAGTTGGAGGCCAGAAACGAAACTGGAGAGCAGCTGGAG 2640
QY 2641 AAGATCAGCCACCAAGACCAACAGTGAACAAAGATCGGCTGTGAACTGAGACAAAGATTG 2700
| | | | |
Db 2641 AAGATCAGCCACCAAGACCAACAGTGAACAAAGATCGGCTGTGAACTGAGAGCAAAAGATTG 2700
QY 2701 CGGAGGTCAGTCTTAGAGCAAGAGAGCAAGAACTGAGAGCTCAAGCCGCTCAAGAG 2760
| | | | |
Db 2701 CGGAGGTCAGTCTTAGAGCAAGAGAGCAAGAACTGAGAGCTCAAGCCGCTCAAGAG 2760
QY 2761 CTACAGGCTCCCTGCGAGAGCGGAGTCAAGTTGAACAGCCCTGAGGCTGACCGGAG 2820
| | | | |
Db 2761 CTACAGGCTCCCTGCGAGAGCGGAGTCAAGTTGAACAGCCCTGAGGCTGACCGGAG 2820
QY 2821 GCCCTGAGAGCCAGCTTCGCGAGCGGAGAGCAGAGCTGGAAGAGACCAACAGCAAGAGCT 2880
| | | | |
Db 2821 GCCCTGAGAGCCAGCTTCGCGAGCGGAGAGCAGAGCTGGAAGAGACCAACAGCAAGAGCT 2880
QY 2881 GAAGAGAGATCCAGGCACTCAAGGCAATGAGATGAATCCAGCCCAATTTGATGCT 2940
| | | | |
Db 2881 GAAGAGAGATCCAGGCACTCAAGGCAATGAGATGAATCCAGCCCAATTTGATGCT 2940
QY 2941 CTTGCTAAGCTGTACTGTATCAAGACCTGGAGAGCAGCTAAACAGCTGACCGAG 3000
| | | | |
Db 2941 CTTGCTAAGCTGTACTGTATCAAGACCTGGAGAGCAGCTAAACAGCTGACCGAG 3000
QY 3001 GACAACTGAACTCAACAAACCAAACTTCTTGTCCAAACAACTGATGAGGCTTCT 3060
| | | | |
Db 3001 GACAACTGAACTCAACAAACCAAACTTCTTGTCCAAACAACTGATGAGGCTTCT 3060
QY 3061 GCGCGCCACGACGAGATTGTACAACTGCGAAGTGAAGTGAACATCTCCGCGGAGATC 3120
| | | | |
Db 3061 GCGCGCCACGACGAGATTGTACAACTGCGAAGTGAAGTGAACATCTCCGCGGAGATC 3120
QY 3061 GCGCGCCACGACGAGATTGTACAACTGCGAAGTGAAGTGAACATCTCCGCGGAGATC 3120
| | | | |
Db 3061 GCGCGCCACGACGAGATTGTACAACTGCGAAGTGAAGTGAACATCTCCGCGGAGATC 3120
QY 3121 ACGGAACGAGAGATGAGCTTACAGGCAAGAGCAACGATGAGGCTTGAAGAACAGC 3180
| | | | |
Db 3121 ACGGAACGAGAGATGAGCTTACAGGCAAGAGCAACGATGAGGCTTGAAGAACAGC 3180
QY 3181 TGCAACATGCTGAGAGAAAGGTCAATGGAATTTGAGGCGCTTAACGATGAGCTGAGAA 3240
| | | | |
Db 3181 TGCAACATGCTGAGAGAAAGGTCAATGGAATTTGAGGCGCTTAACGATGAGCTGAGAA 3240
QY 3241 AAAGAGCGGAGAGGAGGCTGAGAGAGCGTCTGAGTGAAGAAATCCAGTTTGAAG 3300
| | | | |
Db 3241 AAAGAGCGGAGAGGAGGCTGAGAGAGCGTCTGAGTGAAGAAATCCAGTTTGAAG 3300
QY 3301 TGTGCGGTTGAGAGCTGAGAGAGTGTGAGACCGAGAAACGAGACAGAGGCGAGAGCC 3360
| | | | |
Db 3301 TGTGCGGTTGAGAGCTGAGAGAGTGTGAGACCGAGAAACGAGACAGAGGCGAGAGCC 3360

QY 3361 GATCAGGAGATCACCGAGTCTCCGCAAGTGTGAGAGCTGGCAGTGAAGACCAAGGCT 3420
| | | | |
Db 3361 GATCAGGAGATCACCGAGTCTCCGCAAGTGTGAGAGCTGGCAGTGAAGACCAAGGCT 3420
QY 3421 GAGATTCTGCTGCGAGCAGAGGCTCTCAAGAGAGCAAGAGCTGAAGGCGGAGGCTCT 3480
| | | | |
Db 3421 GAGATTCTGCTGCGAGCAGAGGCTCTCAAGAGAGCAAGAGCTGAAGGCGGAGGCTCT 3480
QY 3481 GACAGCTCAATGACCTGAGAGAAAGCATATGCTTGAATGAATGATGCCGAGCTTA 3540
| | | | |
Db 3481 GACAGCTCAATGACCTGAGAGAAAGCATATGCTTGAATGAATGATGCCGAGCTTA 3540
QY 3541 CAGCAGAGCTGAGAGCTGAACGAGAGCTCAACAGAGGCTTCTGGAAGAGCAAGCCAA 3600
| | | | |
Db 3541 CAGCAGAGCTGAGAGCTGAACGAGAGCTCAACAGAGGCTTCTGGAAGAGCAAGCCAA 3600
QY 3601 TTACAGAGAGATGAGCTGCGAGAAATACATTTTCCGTGACTCAAGAGCTGCA 3660
| | | | |
Db 3601 TTACAGAGAGATGAGCTGCGAGAAATACATTTTCCGTGACTCAAGAGCTGCA 3660
QY 3661 GAACTCTAGATGCGGCTGATCTTACGAGACAGAAAGAGTGAAGTGAAGTATCAAGCTG 3720
| | | | |
Db 3661 GAACTCTAGATGCGGCTGATCTTACGAGACAGAAAGAGTGAAGTGAAGTATCAAGCTG 3720
QY 3721 GAAAACATTCAGGTTCTTATTTCTCATGAAGAGTGAAGTGAAGGCACTATTTCTCA 3780
| | | | |
Db 3721 GAAAACATTCAGGTTCTTATTTCTCATGAAGAGTGAAGTGAAGGCACTATTTCTCA 3780
QY 3781 CAAAACCAATTCATGATTTTCTGCAAGCCCAATGGAACCACTGCTAAGAAAGAAAG 3840
| | | | |
Db 3781 CAAAACCAATTCATGATTTTCTGCAAGCCCAATGGAACCACTGCTAAGAAAGAAAG 3840
QY 3841 GTTCTCTGAGTACATGAGCTGAAGCTGAGGCTGAGAGAGAGAAAGCTCGCTGCA 3900
| | | | |
Db 3841 GTTCTCTGAGTACATGAGCTGAAGCTGAGGCTGAGAGAGAGAAAGCTCGCTGCA 3900
QY 3901 GAGCTAGAGAGAGGCTTCAAGAGACCCGATGAGAGCTCCGCTGCGCGGAGAGAGCT 3960
| | | | |
Db 3901 GAGCTAGAGAGAGGCTTCAAGAGACCCGATGAGAGCTCCGCTGCGCGGAGAGAGCT 3960
QY 3961 GCCCACCAGCAACCAAGGACCAACCCACCACTCCAGCCAGCCAGCGGAGGCAAGC 4020
| | | | |
Db 3961 GCCCACCAGCAACCAAGGACCAACCCACCACTCCAGCCAGCCAGCGGAGGCAAGC 4020
QY 4021 ATCCGATGCTCCGCAATCGTGGGCTGCGCAGAGACCAAGCCAGTGCATGAGGCTGCTG 4080
| | | | |
Db 4021 ATCCGATGCTCCGCAATCGTGGGCTGCGCAGAGACCAAGCCAGTGCATGAGGCTGCTG 4080
QY 4081 GCCCGGCAATCAGCCGCAAGAAAGAGTCTTCAACTCCAGAGAAATTAAGTCCGAGCTT 4140
| | | | |
Db 4081 GCCCGGCAATCAGCCGCAAGAAAGAGTCTTCAACTCCAGAGAAATTAAGTCCGAGCTT 4140
QY 4141 AAGGAACGATGACCAACCAATTTCTCAACCGATTCAAGTGAAGTGAAGTGAAGGCT 4200
| | | | |
Db 4141 AAGGAACGATGACCAACCAATTTCTCAACCGATTCAAGTGAAGTGAAGTGAAGGCT 4200
QY 4141 AAGGAACGATGACCAACCAATTTCTCAACCGATTCAAGTGAAGTGAAGTGAAGGCT 4200
| | | | |
Db 4141 AAGGAACGATGACCAACCAATTTCTCAACCGATTCAAGTGAAGTGAAGTGAAGGCT 4200
QY 4201 ACAAGTGTGCTGCTGCTGAGTACCGTGAACCTTGAAGGCGGATCAAAATGCTCTC 4260
| | | | |
Db 4201 ACAAGTGTGCTGCTGCTGAGTACCGTGAACCTTGAAGGCGGATCAAAATGCTCTC 4260
QY 4261 GAAATGTCAGTGAATGCTCAACCCCAAGTGTCAAGTGTCTTCAAGCCAGCACTGCGAGCTTG 4320
| | | | |
Db 4261 GAAATGTCAGTGAATGCTCAACCCCAAGTGTCTTCAAGCCAGCACTGCGAGCTTG 4320
QY 4321 CTTGCTGAATATGCTCAACCTTCAAGAGGCTTCTGCGCTGACAAATGAATCTCCCA 4380
| | | | |
Db 4321 CTTGCTGAATATGCTCAACCTTCAAGAGGCTTCTGCGCTGACAAATGAATCTCCCA 4380
QY 4381 GGTCTCAGACCAAGAGCCAGCAGAGCTTCACTGGAAGGAGTGAAGTGAAGTGGCTCC 4440
| | | | |
Db 4381 GGTCTCAGACCAAGAGCCAGCAGAGCTTCACTGGAAGGAGTGAAGTGAAGTGGCTCC 4440
QY 4441 AAGAAATTAACAAACGAGACAGCAAGGCTGGAGACAGAAATTAATTTCTGAGAGGATCA 4500
| | | | |


```
Db 4441 AGGATATACAAAGAGACAGCAAGGCTGGGACAGGAAGTACATTTGCTCTGGAGGATCA 4500
Qy 4501 AAGTCTCATTTATGCAATGAAGCCAGAGAGCTGGAACAAGCCGGTGGAGAAATTT 4560
Db 4501 AAGTCTCATTTATGCAATGAAGCCAGAGAGCTGGAACAAGCCGGTGGAGAAATTT 4560
Qy 4561 GAGCTGGCTTCCGAGCGGGGAGTATATATTCATATGATGTCGCTGGTCTCCGAATC 4620
Db 4561 GAGCTGGCTTCCGAGCGGGGAGTATATTCATATGATGTCGCTGGTCTCCGAATC 4620
Qy 4621 GCAATATACAGCCAGAGAGATGTCCTCATATCTGAAGATGATCTCACCGCACCC 4680
Db 4621 GCAATATACAGCCAGAGAGATGTCCTCATATCTGAAGATGATCTCACCGCACCC 4680
Qy 4681 ACCCTGGCGCCGGAGAGACCTCTTACTTCTGACTCTCCAGCTTCCCTGCAAAACAGCGC 4740
Db 4681 ACCCTGGCGCCGGAGAGACCTCTTACTTCTGACTCTCCAGCTTCCCTGCAAAACAGCGC 4740
Qy 4741 TGGGTACCGCCCTTAATATGATGTCGAGTGGGAGACTTTCTAGGAAAAAGACAGAA 4800
Db 4741 TGGGTACCGCCCTTAATATGATGTCGAGTGGGAGACTTTCTAGGAAAAAGACAGAA 4800
Qy 4801 GCTGATGCTAACTGCTTGGAAAATCCTCTGCTGAAAATGGAAGTATGACCGCTTAGAC 4860
Db 4801 GCTGATGCTAACTGCTTGGAAAATCCTCTGCTGAAAATGGAAGTATGACCGCTTAGAC 4860
Qy 4861 ATGAATCTGACCGTCCCTTCACTGACCAAGTGTGTGGTGGGACCGAGAGAGGCTC 4920
Db 4861 ATGAATCTGACCGTCCCTTCACTGACCAAGTGTGTGGTGGGACCGAGAGAGGCTC 4920
Qy 4921 TACGCGCTGAATGCTTGAAGAACTCCCTTAATCCCATATCTCCAGAAATGGAAGACGCTTC 4980
Db 4921 TACGCGCTGAATGCTTGAAGAACTCCCTTAATCCCATATCTCCAGAAATGGAAGACGCTTC 4980
Qy 4981 CAAATTTATATTTATCAAGGACCTGAGAAAGCTACTATGATAGCAGAGAGAGCGGACA 5040
Db 4981 CAAATTTATATTTATCAAGGACCTGAGAAAGCTACTATGATAGCAGAGAGAGCGGACA 5040
Qy 5041 CTGTGCTTGTGGACCTGAGAAAGTGAACAGTCCCTGGCCCAAGTCCCACTGCTGCC 5100
Db 5041 CTGTGCTTGTGGACCTGAGAAAGTGAACAGTCCCTGGCCCAAGTCCCACTGCTGCC 5100
Qy 5101 CAGCCCGACATCTCAACCCAACTTTTGAAGCTGTCAAGGGCTGCCACTTGTGTGGGCA 5160
Db 5101 CAGCCCGACATCTCAACCCAACTTTTGAAGCTGTCAAGGGCTGCCACTTGTGTGGGCA 5160
Qy 5161 GGCAGATTTAGAGACGGGCTGTGATCTGTGACGACATGCCAGCAAAAGTCTCATTTCTC 5220
Db 5161 GGCAGATTTAGAGACGGGCTGTGATCTGTGACGACATGCCAGCAAAAGTCTCATTTCTC 5220
Qy 5221 CGCTACACGAAACCTCAGCAATATCTGATCCGGAAGAGATAGAGACTTCAGAGCCC 5280
Db 5221 CGCTACACGAAACCTCAGCAATATCTGATCCGGAAGAGATAGAGACTTCAGAGCCC 5280
Qy 5281 TGCAGCTGATCTCACTTCAACAATTAAGTATCTTATGGAACCAATTAATTTTCAGAA 5340
Db 5281 TGCAGCTGATCTCACTTCAACAATTAAGTATCTTATGGAACCAATTAATTTTCAGAA 5340
Qy 5341 ATGCAATATGAGACGATGACGCTCGAGAAATTCCTGGATTAAGATAGCAATTCCTTGGCA 5400
Db 5341 ATGCAATATGAGACGATGACGCTCGAGAAATTCCTGGATTAAGATAGCAATTCCTTGGCA 5400
Qy 5401 CCTGCTGTGTGTTCGCGCTCTTTCACAACAGCTTCCCTGTCTCAATCGTGCAGGTGAACAC 5460
Db 5401 CCTGCTGTGTGTTCGCGCTCTTTCACAACAGCTTCCCTGTCTCAATCGTGCAGGTGAACAC 5460
Qy 5461 GAGAGGACGAGAGAGAGATCTTGTGTTCACAAGATTTGAGTGTTCGTGGATTC 5520
Db 5461 GAGAGGACGAGAGAGAGATCTTGTGTTCACAAGATTTGAGTGTTCGTGGATTC 5520
Qy 5521 TACGGAAGCTAGCGGACAGACGATCTCAAGTGGAGCGCTTACCTTTGGCCTTTGCC 5580
Db 5521 TACGGAAGCTAGCGGACAGACGATCTCAAGTGGAGCGCTTACCTTTGGCCTTTGCC 5580
```

```
Db 5521 TACGGAAGCTAGCGGACAGACGATCTCAAGTGGAGCGCTTACCTTTGGCCTTTGCC 5580
Qy 5581 TACAGAAACCTATCTGTTTGTGACCCACTTCACTCAAGTAATGATGATCAG 5640
Db 5581 TACAGAAACCTATCTGTTTGTGACCCACTTCACTCAAGTAATGATGATCAG 5640
Qy 5641 GCAAGCTCTCAGACAGAGACCCCTGCGCCAGCGTACTGACATTCGGAACCCGCGTAC 5700
Db 5641 GCAAGCTCTCAGACAGAGACCCCTGCGCCAGCGTACTGACATTCGGAACCCGCGTAC 5700
Qy 5701 CTGGGCTCTGCAATTTCTCAGAGAGCATTTACTTGGCGTCTCATATCAGATTAATTA 5760
Db 5701 CTGGGCTCTGCAATTTCTCAGAGAGCATTTACTTGGCGTCTCATATCAGATTAATTA 5760
Qy 5761 AGGTCATTTGCTGCAAGGAAAACCTGTGAGAGATCGGCACTGAACCAACCGGCGC 5820
Db 5761 AGGTCATTTGCTGCAAGGAAAACCTGTGAGAGATCGGCACTGAACCAACCGGCGC 5820
Qy 5821 CCGTCAACCTCCCGAGAGCCCAACAGCGAGGCGCAACCGCATTCACAGAGCATTC 5880
Db 5821 CCGTCAACCTCCCGAGAGCCCAACAGCGAGGCGCAACCGCATTCACAGAGCATTC 5880
Qy 5881 ACCAAGCGGTGGCTCTCAGACCCAGCGCCCGCAAGGCCCGCAAGCCCGAGAGCCA 5940
Db 5881 ACCAAGCGGTGGCTCTCAGACCCAGCGCCCGCAAGGCCCGCAAGCCCGAGAGCCA 5940
Qy 5941 AGCACACCCACCGCTTACCGAGAGGGGCGAATCGAAGTCTCTGCTGCC 6000
Db 5941 AGCACACCCACCGCTTACCGAGAGGGGCGAATCGAAGTCTCTGCTGCC 6000
Qy 6001 CCGCCCTGAGACGAGAGAGTCCCGCGCGCGATCTGACACCGGAGAGAGCGCTCC 6060
Db 6001 CCGCCCTGAGACGAGAGAGTCCCGCGCGCGATCTGACACCGGAGAGAGCGCTCC 6060
Qy 6061 CCGCGAGGCTGTTTGAAGACAGACAGAGGCGGCTCTGCGGAGACCGTGAAGACC 6120
Db 6061 CCGCGAGGCTGTTTGAAGACAGACAGAGGCGGCTCTGCGGAGACCGTGAAGACC 6120
Qy 6121 CCGCTGTCCAGGTGAACAAAGTCTGGAGCAAGTCTTCAATTA 6165
Db 6121 CCGCTGTCCAGGTGAACAAAGTCTGGAGCAAGTCTTCAATTA 6165
RESULT 6
AX671112 6156 bp DNA linear PART 27-MAR-2003
LOCUS AX671112
DEFINITION Sequence 8 from Patent WO03004523.
ACCESSION AX671112
VERSION AX671112.1 GI:29329572
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 Zhu, Z.
Regulation of human citron rho/rac-interacting kinase
Patent: WO 03004523-A 8 16-JAN-2003;
Bayer Aktiengesellschaft (DE)
FEATURES
source 1..6156
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 99.6%; Score 6142.2; DB 6; Length 6156;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 6147; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
Qy 1 ATGTTGAAGTTCAATATGAGACGCGGAATCCTTTGAGATGCTGTGCTGAACCCATT 60
Db 1 ATGTTGAAGTTCAATATGAGACGCGGAATCCTTTGAGATGCTGTGCTGAACCCATT 60
```


QY 61 GCCAGCCGGGCTCCAGGCTGAATCTGTTCACAGGGGAAACACCCCTTTATGACTCAA 120
DB 61 GCCAGCCGGGCTCCAGGCTGAATCTGTTCACAGGGGAAACACCCCTTTATGACTCAA 120
QY 121 CAGCAATGTCCTCTCTTTCCCGAAGAGATATTAATGTCCTCTTTGTTCTTTGAA 180
DB 121 CAGCAATGTCCTCTCTTTCCCGAAGAGATATTAATGTCCTCTTTGTTCTTTGAA 180
QY 181 GAATGCAAGCTGCTCTGATGAAGATTAAAGCACTTGAGCACTTTGTCCGGAAGTAT 240
DB 181 GAATGCAAGCTGCTCTGATGAAGATTAAAGCACTTGAGCACTTTGTCCGGAAGTAT 240
QY 241 TCCGACACCACTAGCTGATGACAGAGACTCCAGCTTCGCGAAAGACTTGAAGTACA 300
DB 241 TCCGACACCACTAGCTGATGACAGAGACTCCAGCTTCGCGAAAGACTTGAAGTACA 300
QY 301 AGCTCTGAGAGTTGTGTCTCTTTGCTGAAGTCAAGGTGTTAAGAGAAAGCAACCGG 360
DB 301 AGCTCTGAGAGTTGTGTCTCTTTGCTGAAGTCAAGGTGTTAAGAGAAAGCAACCGG 360
QY 361 GACATCTATCTATGAAGATGAGAGAAAGGCTTTATTTGGCCGAGAGGTTTCA 420
DB 361 GACATCTATCTATGAAGATGAGAGAAAGGCTTTATTTGGCCGAGAGGTTTCA 420
QY 421 TTTTGGAGGAGAGCGGAAATATTTATCTGAAGCAAGCCGTTGATCCCCCAATTA 480
DB 421 TTTTGGAGGAGAGCGGAAATATTTATCTGAAGCAAGCCGTTGATCCCCCAATTA 480
QY 481 CAGTATGCTTTTGAAGCAAAAAATCACCTTTATCTGCTCATGGAATATCACCTGAGGG 540
DB 481 CAGTATGCTTTTGAAGCAAAAAATCACCTTTATCTGCTCATGGAATATCACCTGAGGG 540
QY 541 GACTGTGTCATCTTTGAATAGATAGAGACAGTTAGATGAAACCTGATACGTTT 600
DB 541 GACTGTGTCATCTTTGAATAGATAGAGACAGTTAGATGAAACCTGATACGTTT 600
QY 601 TACCTAGCTGAGCTGATTTTGGCTGTTCAAGAGCTTCATCTGATGAGGATACGTGATCA 660
DB 601 TACCTAGCTGAGCTGATTTTGGCTGTTCAAGAGCTTCATCTGATGAGGATACGTGATCA 660
QY 661 GACATCAAGCTGAGAACTTCTGCTTGAACCGCAACAGAACTCAAGCTGAGGATTTT 720
DB 661 GACATCAAGCTGAGAACTTCTGCTTGAACCGCAACAGAACTCAAGCTGAGGATTTT 720
QY 721 GGAATCGCGGAGAAATGAATTCAAACAGATGTTGAATGCCAATCCCGATTTGGGACC 780
DB 721 GGAATCGCGGAGAAATGAATTCAAACAGATGTTGAATGCCAATCCCGATTTGGGACC 780
QY 781 CCAGATTTACATGAGCTCTGAAAGTCTGATCTGATGAACGGGATGAGAAAGGCACTTAC 840
DB 781 CCAGATTTACATGAGCTCTGAAAGTCTGATCTGATGAACGGGATGAGAAAGGCACTTAC 840
QY 841 GGCCTGAGCTGTAATGCTGAGTCAAGTGGCGTGAATTCCTATGATGATTTATGGAGA 900
DB 841 GGCCTGAGCTGTAATGCTGAGTCAAGTGGCGTGAATTCCTATGATGATTTATGGAGA 900
QY 901 TCCCTGAGCTGTAATGCTGAGTCAAGTGGCGTGAATTCCTATGATGATTTATGGAGA 960
DB 901 TCCCTGAGCTGTAATGCTGAGTCAAGTGGCGTGAATTCCTATGATGATTTATGGAGA 960
QY 961 TTTTGAATTTTCCAGATGACCCCAAGTGAAGTGAAGTCTTTGATCTGATTCAAAGC 1020
DB 961 TTTTGAATTTTCCAGATGACCCCAAGTGAAGTGAAGTCTTTGATCTGATTCAAAGC 1020
QY 1021 TTGTGTGCGGCGAAGAAAGAGACTGAAGTTTGAAGGCTTTGTCTGCACTCTTTCTTC 1080
DB 1021 TTGTGTGCGGCGAAGAAAGAGACTGAAGTTTGAAGGCTTTGTCTGCACTCTTTCTTC 1080
QY 1081 TCTTAAATTTGAATGAGCAATCTTCTGTAATCTCTCTCTCTCTCTCTCTCTCTCAAG 1140
DB 1081 TCTTAAATTTGAATGAGCAATCTTCTGTAATCTCTCTCTCTCTCTCTCTCTCTCAAG 1140

QY 1141 TCTGAGATGACACCTTCCAAATTTTGAATGAACGAGAAAGAAATTCGTGGTTTGCCTCT 1200
DB 1141 TCTGAGATGACACCTTCCAAATTTTGAATGAACGAGAAAGAAATTCGTGGTTTGCCTCT 1200
QY 1201 CGGTGCAAGCTGAGCCCTCAGGCTTCTGGGGGAAAGAACTGGCGTTTGTGGGGTTTCG 1260
DB 1201 CGGTGCAAGCTGAGCCCTCAGGCTTCTGGGGGAAAGAACTGGCGTTTGTGGGGTTTCG 1260
QY 1261 TACAGCAAGCACTGGGGATTTCTGGTGAATCTGAAGTCTGTGTGTGTGTGTGTGTGTGT 1320
DB 1261 TACAGCAAGCACTGGGGATTTCTGGTGAATCTGAAGTCTGTGTGTGTGTGTGTGTGTGT 1320
QY 1321 CTTGCAAGCACTGCTCAGTGAAGAAAGAACTTCTCATCAAAAGCAAGGCTTCAAGAC 1380
DB 1321 CTTGCAAGCACTGCTCAGTGAAGAAAGAACTTCTCATCAAAAGCAAGGCTTCAAGAC 1380
QY 1381 TCTCAGAGCAAGTGTACAAAGATGAGAGGAAATGACCGGTTTACATCGGAGAGTCA 1440
DB 1381 TCTCAGAGCAAGTGTACAAAGATGAGAGGAAATGACCGGTTTACATCGGAGAGTCA 1440
QY 1441 GAGTGAAGCTGTGCTTATGTCAGAAAGAGTGAAGCTTGAAGCTTGAAGCTTGAAGCT 1500
DB 1441 GAGTGAAGCTGTGCTTATGTCAGAAAGAGTGAAGCTTGAAGCTTGAAGCTTGAAGCT 1500
QY 1501 TCCCTCTGAGAGAGACCTTGTCTTACCTACATCAAGATGCAATGCAATGCAATGCAAT 1560
DB 1501 TCCCTCTGAGAGAGACCTTGTCTTACCTACATCAAGATGCAATGCAATGCAATGCAAT 1560
QY 1561 TTGGAGCAAGCAAGATGAGAGTGTCCAGAGAGATGACAAACATCTGAGCTTCTCAT 1620
DB 1561 TTGGAGCAAGCAAGATGAGAGTGTCCAGAGAGATGACAAACATCTGAGCTTCTCAT 1620
QY 1621 GATATCAGAGAGAGAGCCGAGAGCTCCAAAGAAATCAAGAGAGAGTACAGGCTCAA 1680
DB 1621 GATATCAGAGAGAGAGCCGAGAGCTCCAAAGAAATCAAGAGAGAGTACAGGCTCAA 1680
QY 1681 GTGGAAGAAATGAGTGTGATGATGAATCACTTGAAGAGATCTTGTCTCAGCAAGAA 1740
DB 1681 GTGGAAGAAATGAGTGTGATGATGAATCACTTGAAGAGATCTTGTCTCAGCAAGAA 1740
QY 1741 CGAGTGAATCTTACGAATTTGAGCTGAGAGATCTTGGCTTCTGCTGAAGAAATTCAG 1800
DB 1741 CGAGTGAATCTTACGAATTTGAGCTGAGAGATCTTGGCTTCTGCTGAAGAAATTCAG 1800
QY 1801 CGGAAAGCAGAGAAATGTCAGATTAATCTTGAAGCTTAAGATCAAGGAAAGCTGAA 1860
DB 1801 CGGAAAGCAGAGAAATGTCAGATTAATCTTGAAGCTTAAGATCAAGGAAAGCTGAA 1860
QY 1861 GTGGAAGAAATGAGTGTGATGATGAATCACTTGAAGAGATCTTGTCTCAGCAAGAA 1920
DB 1861 GTGGAAGAAATGAGTGTGATGATGAATCACTTGAAGAGATCTTGTCTCAGCAAGAA 1920
QY 1921 CTCGAAGAAATCTGAGAAAGCTGTTAAAGCCAGACCGAGGCTCACCGAGCTGCTGAG 1980
DB 1921 CTCGAAGAAATCTGAGAAAGCTGTTAAAGCCAGACCGAGGCTCACCGAGCTGCTGAG 1980
QY 1981 AATATCGCGAGCAAGAGAGAGCGAGAGGAGGAGTGAAGAGCTGAGAAAGCCGAGAG 2040
DB 1981 AATATCGCGAGCAAGAGAGAGCGAGAGGAGGAGTGAAGAGCTGAGAAAGCCGAGAG 2040
QY 2041 GATTTCTTGAAGAGCTCAGAAAGAAAGCTGTGAAAGCTGAGAAAGCCGCAATTCCTG 2100
DB 2041 GATTTCTTGAAGAGCTCAGAAAGAAAGCTGTGAAAGCTGAGAAAGCCGCAATTCCTG 2100
QY 2101 GAGAAAGAGTAAAGAGACTGAGAGCACTGAGAGCTTGAAGAAAGAGTGAAGTGAAG 2160
DB 2101 GAGAAAGAGTAAAGAGACTGAGAGCACTGAGAGCTTGAAGAAAGAGTGAAGTGAAG 2160
QY 2161 ATCCAGCAAAATCCCAAGATTCAGAGATGAGTGAAGTGAAGTGAAGTGAAGTGAAG 2220
DB 2161 ATCCAGCAAAATCCCAAGATTCAGAGATGAGTGAAGTGAAGTGAAGTGAAGTGAAG 2220
QY 2221 AAACATCGGAGGCGCAAGTCTCAGGCTCAGCACTGAAGTGAAGTGAAGTGAAGTGAAG 2280

|||||
Db 2221 AAACATCGGGAGGCCCAAGTCTCAGCCAGACCTTAGAAGTGCACCTGAAACAGAAAGAG 2280
|||
Qy 2281 CAGACCTATGAGAAAAAGATTAAAGTTGGACAATCAGATAAAGAAAGACTGCGTGAC 2340
|||
Db 2281 CAGACCTATGAGAAAAAGATTAAAGTTGGACAATCAGATAAAGAAAGACTGCGTGAC 2340
|||
Qy 2341 AAGGAGACCTGGAGAACATGATGACAGACACGAGAGAGAGGCCCATGAGAAAGGGGAAA 2400
|||
Db 2341 AAGGAGACCTGGAGAACATGATGACAGACACGAGAGAGAGGCCCATGAGAAAGGGGAAA 2400
|||
Qy 2401 ATTCTCAGCAACAGAAAGCGCATGATCATGCTATGATTTCAAGATCAGATCCCTGGAA 2460
|||
Db 2401 ATTCTCAGCAACAGAAAGCGCATGATCATGCTATGATTTCAAGATCAGATCCCTGGAA 2460
|||
Qy 2461 CAGAGATTGTGGAACTGTCTGAAGCCAAATPAACTTCACAAATAGCAGCTTTTAAAC 2520
|||
Db 2461 CAGAGATTGTGGAACTGTCTGAAGCCAAATPAACTTCACAAATAGCAGCTTTTAAAC 2520
|||
Qy 2521 CAAGAAGACATGAAGGCCCAAGAAAGATGATTTCTGAATCTCAGGCAACAGAAATTTTAC 2580
|||
Db 2521 CAAGAAGACATGAAGGCCCAAGAAAGATGATTTCTGAATCTCAGGCAACAGAAATTTTAC 2580
|||
Qy 2581 CTGAGACACAGGCTGGGAAGTTGAGAGCCCAAGAACGAAAACTGAGAGCAGCTGGAG 2640
|||
Db 2581 CTGAGACACAGGCTGGGAAGTTGAGAGCCCAAGAACGAAAACTGAGAGCAGCTGGAG 2640
|||
Qy 2641 AAGATCAGCCACCAAGACCAACAGTACAGAAATCGCTGCTGGAACTGAGACAAAGATTG 2700
|||
Db 2641 AAGATCAGCCACCAAGACCAACAGTACAGAAATCGCTGCTGGAACTGAGACAAAGATTG 2700
|||
Qy 2701 CGGAGAGTCAGTCTAGAGCAGAGAGCAGAGAGAACTGGAGCTCAAGGCCAGCTCACAAG 2760
|||
Db 2701 CGGAGAGTCAGTCTAGAGCAGAGAGCAGAGAGAACTGGAGCTCAAGGCCAGCTCACAAG 2760
|||
Qy 2761 CTACAGCTCTCCCTGAGAGAGCGCAGATCAAGCTTGAACAGCCCTGAGAGCTGACCGGCG 2820
|||
Db 2761 CTACAGCTCTCCCTGAGAGAGCGCAGATCAAGCTTGAACAGCCCTGAGAGCTGACCGGCG 2820
|||
Qy 2821 GCCCTGAGAGCCAGCTTGCAGGCGAGAGACAGAGCTGAGAGACACACAGCAAGACT 2880
|||
Db 2821 GCCCTGAGAGCCAGCTTGCAGGCGAGAGACAGAGCTGAGAGACACACAGCAAGACT 2880
|||
Qy 2881 GAAGAGAGATCCAGGCACTCAGGCACATAGAGATGAATCCAGGCGAAATTTGATGCT 2940
|||
Db 2881 GAAGAGAGATCCAGGCACTCAGGCACATAGAGATGAATCCAGGCGAAATTTGATGCT 2940
|||
Qy 2941 CTTCTGAACAGCTGTACTGTAAATCAGACCTGAGAGAGAGCTAAACAGCTGACCGAG 3000
|||
Db 2941 CTTCTGAACAGCTGTACTGTAAATCAGACCTGAGAGAGAGCTAAACAGCTGACCGAG 3000
|||
Qy 3001 GACACGCTGAATCTCAACCAAACTTCTGACAACTCGATGAGGCTTCT 3060
|||
Db 3001 GACACGCTGAATCTCAACCAAACTTCTGACAACTCGATGAGGCTTCT 3060
|||
Qy 3061 GCGCGCAACGAGAGATTGTAACTGCGAGAGTGAAGTGAACCATTCGCGCGGAGATC 3120
|||
Db 3061 GCGCGCAACGAGAGATTGTAACTGCGAGAGTGAAGTGAACCATTCGCGCGGAGATC 3120
|||
Qy 3121 ACGGAACGAGATGACGCTTACAGCCAGAAACAAACGATGAGGCTCTGAAGACACG 3180
|||
Db 3121 ACGGAACGAGATGACGCTTACAGCCAGAAACAAACGATGAGGCTCTGAAGACACG 3180
|||
Qy 3181 TGCACCATGCTGAGAGAACAGTCAATGATTTGAGAGCCCTAAACGATGAGCTGTAGAA 3240
|||
Db 3181 TGCACCATGCTGAGAGAACAGTCAATGATTTGAGAGCCCTAAACGATGAGCTGTAGAA 3240
|||
Qy 3241 AAAGAGCGGCACTGGAGAGCCCTGAGAGAGGCTCTGAGTGAAGAAATCCAGTTTGA 3300
|||
Db 3241 AAAGAGCGGCACTGGAGAGCCCTGAGAGAGGCTCTGAGTGAAGAAATCCAGTTTGA 3300
|||
Qy 3301 TGTCCGGTTTCAGAGCTGAGAGATGCTGACACCGAGAAACAGAGCAGGCGAGAGCC 3360
|||

|||||
Db 3301 TGTCCGGTTTCAGAGCTGAGAGATGCTGACACCGAAACAGAGCAGGCGAGAGCC 3360
|||
Qy 3361 GATCAGCGGATCACCGAGTCTCCCGAGGTGTGAGAGCTGACAGTGAAGAGACCAAGGCT 3420
|||
Db 3361 GATCAGCGGATCACCGAGTCTCCCGAGGTGTGAGAGCTGACAGTGAAGAGACCAAGGCT 3420
|||
Qy 3421 GAGATTCTCGCTCTGAGAGCGGCTCTCAAGAGCAGAAAGCTGAAGGCTCGAGAGCTTCT 3480
|||
Db 3421 GAGATTCTCGCTCTGAGAGCGGCTCTCAAGAGCAGAAAGCTGAAGGCTCGAGAGCTTCT 3480
|||
Qy 3481 GACAACTCATGACCTGAGAGAAAGCATGCTATGCTTGAATGAATGCTCCGAAAGCTTA 3540
|||
Db 3481 GACAACTCATGACCTGAGAGAAAGCATGCTATGCTTGAATGAATGCTCCGAAAGCTTA 3540
|||
Qy 3541 CAGCAGAGCTGAGAGCTGAACGAGAGCTTCAACAGAGCTTCTGAGAGACCAAGCCAAA 3600
|||
Db 3541 CAGCAGAGCTGAGAGCTGAACGAGAGCTTCAACAGAGCTTCTGAGAGACCAAGCCAAA 3600
|||
Qy 3601 TTACAGCAGAGATGAGACCTGCAAGAAATCAATTTTCCGTCTGACTCAGAGACTGCAA 3660
|||
Db 3601 TTACAGCAGAGATGAGACCTGCAAGAAATCAATTTTCCGTCTGACTCAGAGACTGCAA 3660
|||
Qy 3661 GAAGCTCTAGATCGGCTGATCTACTGAGAGACAGAAAGAGTGAATTTGAGATACGCTG 3720
|||
Db 3661 GAAGCTCTAGATCGGCTGATCTACTGAGAGACAGAAAGAGTGAATTTGAGATACGCTG 3720
|||
Qy 3721 GAAACATTCAGGTTCTATTTCTCATGAAGAGTGAAGAGCACTATTTCTCAA 3780
|||
Db 3721 GAAACATTCAGGTTCTATTTCTCATGAAGAGTGAAGAGCACTATTTCTCAA 3780
|||
Qy 3781 CAACCAAACTCATGATTTTCTGCAAGCCAAATGCAACCTCTTAATAAAGAAAG 3840
|||
Db 3781 CAACCAAACTCATGATTTTCTGCAAGCCAAATGCAACCTCTTAATAAAGAAAG 3840
|||
Qy 3841 GTTCTCTCAGTACATGAGTGAAGCTGAGCTGAGAGAGAGAACTGCTGTGCA 3900
|||
Db 3841 GTTCTCTCAGTACATGAGTGAAGCTGAGCTGAGAGAGAGAACTGCTGTGCA 3900
|||
Qy 3901 GAGCTAGAGAAAGCCCTTCAAGAAAGCCCGCATGAGACTCGGCTCCGCGGAGAGAGCT 3960
|||
Db 3901 GAGCTAGAGAAAGCCCTTCAAGAAAGCCCGCATGAGACTCGGCTCCGCGGAGAGAGCT 3960
|||
Qy 3961 GCCCACCGGAAGCAACGGAACCAACCAACCCATCCAGCCAGCCACCGCGAGGCGAGCAG 4020
|||
Db 3961 GCCCACCGGAAGCAACGGAACCAACCAACCCATCCAGCCAGCCACCGCGAGGCGAGCAG 4020
|||
Qy 4021 ATGCGCATGTCGCGCATGCGGAGTGCAGAGAGACACAGCCAGTGCATGAGGCTGCTG 4080
|||
Db 4021 ATGCGCATGTCGCGCATGCGGAGTGCAGAGAGACACAGCCAGTGCATGAGGCTGCTG 4080
|||
Qy 4081 GCCCGGCAATCCAGCCGCGAGAAAGAGTCTCAACTCCAGAGAAATTTAGTGGCGTCTT 4140
|||
Db 4081 GCCCGGCAATCCAGCCGCGAGAAAGAGTCTCAACTCCAGAGAAATTTAGTGGCGTCTT 4140
|||
Qy 4141 AAGGAACGATGACACCAAAATTTCTCACCAGATTCAGAGTGAAGTGAACATGAGAGCC 4200
|||
Db 4141 AAGGAACGATGACACCAAAATTTCTCACCAGATTCAGAGTGAAGTGAACATGAGAGCC 4200
|||
Qy 4201 ACAAAAGTGTGTGTGTGTGTGATACCGTGCATTTTGAACGCGCAGAGCAATCCAAATGTCTC 4260
|||
Db 4201 ACAAAAGTGTGTGTGTGTGTGATACCGTGCATTTTGAACGCGCAGAGCAATCCAAATGTCTC 4260
|||
Qy 4261 GAATGTCAAGTGAATGTTCACCCCAAGTCTCAAGTGTCTTCCAGCACTTCCGCGCTTG 4320
|||
Db 4261 GAATGTCAAGTGAATGTTCACCCCAAGTCTCAAGTGTCTTCCAGCACTTCCGCGCTTG 4320
|||
Qy 4321 CTTGCTGAATATGCAACACACTTCAACGAGAGCTTCTGCGTGAACAAATGAATCTCCCA 4380
|||
Db 4321 CTTGCTGAATATGCAACACACTTCAACGAGAGCTTCTGCGTGAACAAATGAATCTCCCA 4380
|||
Qy 4381 GGTCTCCAGACCAAGAGCCCAAGAGAGCTTTCAGCTGGAAGGTTGAAGGTTGCCC 4440
|||
Db 4381 GGTCTCCAGACCAAGAGCCCAAGAGAGCTTTCAGCTGGAAGGTTGAAGGTTGCCC 4440
|||

QY 1 ATGTTGAAGTTCAAAATATGAGGCGGGAATCCTTTGATGCTGGTGCCTGAACCAATT 60
Db 1 ATGTTGAAGTTCAAAATATGAGGCGGGAATCCTTTGATGCTGGTGCCTGAACCAATT 60
QY 61 GCCAGCGGGGCTTCCAGGCTGAATCTGTTCTTCCAGGGGAAACCAACCTTATGACTCAA 120
Db 61 GCCAGCGGGGCTTCCAGGCTGAATCTGTTCTTCCAGGGGAAACCAACCTTATGACTCAA 120
QY 121 CAGCAATGTTCTCTCTTTTCCCGAAGGGAATTAATGATGCTCTTTGTTCTTTGAA 180
Db 121 CAGCAATGTTCTCTCTTTTCCCGAAGGGAATTAATGATGCTCTTTGTTCTTTGAA 180
QY 181 GAATGCACTGAGCTGCTGATGAAGATTAAAGCACTGAAGCACTTGTCCGGA---G 237
Db 181 GAATGCACTGAGCTGCTGATGAAGATTAAAGCACTGAAGCACTTGTCCGGAAGTG 240
QY 238 TATTCGACACCAATAGCTGATTACAGAGACTCCAGCTTCGCGAAGAAGACTTGAAGTC 297
Db 241 TATTCGACACCAATAGCTGATTACAGAGACTCCAGCTTCGCGAAGAAGACTTGAAGTC 300
QY 298 AGAAGTCTTGAGTTGGTCACTTTGCTGAAGTGCAGGTGGTAAAGAGAAAGCAACC 357
Db 301 AGAAGTCTTGAGTTGGTCACTTTGCTGAAGTGCAGGTGGTAAAGAGAAAGCAACC 360
QY 358 GGGGACATCTATGCTAATGAAGATGAAGAAGAGCTTATTTGGCCGAGAGCAAGTT 417
Db 361 GGGGACATCTATGCTAATGAAGATGAAGAAGAGCTTATTTGGCCGAGAGCAAGTT 420
QY 418 TCATTTTGTGAGAGAAGCGGAACATTTATCTCGAAGCACAAGCCGCTGATCCCCAA 477
Db 421 TCATTTTGTGAGAGAAGCGGAACATTTATCTCGAAGCACAAGCCGCTGATCCCCAA 480
QY 478 TTACAGTATGCTTTTCAAGCAAAAATCACCTTTATCTGGTCAATGGAATACAGCTTGA 537
Db 481 TTACAGTATGCTTTTCAAGCAAAAATCACCTTTATCTGGTCAATGGAATACAGCTTGA 540
QY 538 GGGGACTTGCTGCACTTTGAATAGATAGAGAACAGTTAGATGAAGAACTGATACAG 597
Db 541 GGGGACTTGCTGCACTTTGAATAGATAGAGAACAGTTAGATGAAGAACTGATACAG 600
QY 598 TTTTACCTTGAAGTGAATTTTGGCTGTTCAAGCGTTCACTGAATGGGATAGTGCAAT 657
Db 601 TTTTACCTTGAAGTGAATTTTGGCTGTTCAAGCGTTCACTGAATGGGATAGTGCAAT 660
QY 658 CGAGACATCAAGCTGAGAAATTTCTGTTGACCGACAGAGACATCAAGCTGTTGAT 717
Db 661 CGAGACATCAAGCTGAGAAATTTCTGTTGACCGACAGAGACATCAAGCTGTTGAT 720
QY 718 TTTGATCTGCGCGGAAAATGAATTCAAACAAGATGTAATGCAATGCCAACTCCCGATTGG 777
Db 721 TTTGATCTGCGCGGAAAATGAATTCAAACAAGATGTAATGCAATGCCAACTCCCGATTGG 780
QY 778 ACCCCAGATTACATGAGCTCTGTAAGTGTGACTGTGATGAACGGGATGGAAAAAGCAC 837
Db 781 ACCCCAGATTACATGAGCTCTGTAAGTGTGACTGTGATGAACGGGATGGAAAAAGCAC 840
QY 838 TACGCGCTGGAAGTGTGATGCTGTTGATGAGTGGCGTGAATTTATGAG 897
Db 841 TACGCGCTGGAAGTGTGATGCTGTTGATGAGTGGCGTGAATTTATGAG 900
QY 898 AATGCCCCCTTGGAGAGGGAACCTCTGCAAGACCTTGAATTAACCTTATGATTTCCAG 957
Db 901 AATGCCCCCTTGGAGAGGGAACCTCTGCAAGACCTTGAATTAACCTTATGATTTCCAG 960
QY 958 CGGTTTTGAATTTTCAGATGACCCCAAGTGAAGTGAAGTCTTTCTGATCTGAATCAA 1017
Db 961 CGGTTTTGAATTTTCAGATGACCCCAAGTGAAGTGAAGTCTTTCTGATCTGAATCAA 1020
QY 1018 AGCTTGTGTGCGGCAAGAAAGAGACTGAAGTTGAAGTCTTTGCTGCCATCTTTTC 1077
Db 1021 AGCTTGTGTGCGGCAAGAAAGAGACTGAAGTTGAAGTCTTTGCTGCCATCTTTTC 1080
QY 1078 TTCTCTAAATATGACTGGAACAACATTCGTAACTCTCTCCCTCCCTTGGTCTCCACCTTC 1137

Db 1081 TTCTCTAAATATGACTGGAACAACATTCGTAACTCTCTCCCTCCCTTGGTCTCCACCTTC 1140
QY 1138 AAGTGTACGATGACACCTTCCAAATTTTGAATGAACAGAGAAAGATTGTGGGTTTCATCC 1197
Db 1141 AAGTGTACGATGACACCTTCCAAATTTTGAATGAACAGAGAAAGATTGTGGGTTTCATCC 1200
QY 1198 TCTTCGTTGCAAGTGAAGCTTCAAGCTTCTGCGGTGAAGAACTCCGTTTGTGGGGTTT 1257
Db 1201 TCTTCGTTGCAAGTGAAGCTTCAAGCTTCTGCGGTGAAGAACTCCGTTTGTGGGGTTT 1260
QY 1258 TCGTACAGCAAGGCACTGGGGATTCTTGTGAATCTGAAGCTGTGTTGTGGGTGGAG 1317
Db 1261 TCGTACAGCAAGGCACTGGGGATTCTTGTGAATCTGAAGCTGTGTTGTGGGTGGAG 1320
QY 1318 TCCCTTCCCAAGACTAGCTCATAGAAAAGAACTTCTCATCAAAAGCAAGAGCTACAA 1377
Db 1321 TCCCTTCCCAAGACTAGCTCATAGAAAAGAACTTCTCATCAAAAGCAAGAGCTACAA 1380
QY 1378 GACTCTCAGGCAAGTGTCACAAGATGAGCAGAGAAATGACCCGGTTACATCGAGAGTG 1437
Db 1381 GACTCTCAGGCAAGTGTCACAAGATGAGCAGAGAAATGACCCGGTTACATCGAGAGTG 1440
QY 1438 TCAGAGTGAAGCTGTGCTTATGCTCAAGAGAGTGAAGCTGAAGCTTGAAGTCTCAG 1497
Db 1441 TCAGAGTGAAGCTGTGCTTATGCTCAAGAGAGTGAAGCTGAAGCTTGAAGTCTCAG 1500
QY 1498 AGATCCCTCTCGAGAGAGACCTTGTCTACCTACATCAAGAAATGCAAGTACCTTAAGCGA 1557
Db 1501 AGATCCCTCTCGAGAGAGACCTTGTCTACCTACATCAAGAAATGCAAGTACCTTAAGCGA 1560
QY 1558 AGTTTGAAGCAAGCAGATGAGAGTGTCCAGAGAGATGACAAAGCACTGCATCTTCTC 1617
Db 1561 AGTTTGAAGCAAGCAGATGAGAGTGTCCAGAGAGATGACAAAGCACTGCATCTTCTC 1620
QY 1618 CATGATATCAGAGAGAGAGCGGGAAGCTTCCAAAGAAATCAAGAGAGAGATTCAGAGCT 1677
Db 1621 CATGATATCAGAGAGAGAGCGGGAAGCTTCCAAAGAAATCAAGAGAGAGATTCAGAGCT 1680
QY 1678 CAAAGTGAAGAAATGAGTTGATGATCAAGTTGGAAGAGATCTTGTCTCAGCAAGA 1737
Db 1681 CAAAGTGAAGAAATGAGTTGATGATCAAGTTGGAAGAGATCTTGTCTCAGCAAGA 1740
QY 1738 AGACGAGTGAATCTTACGAATCTGAGCTGAGAGACTCTCGGCTTGCTGTAAGAAATTC 1797
Db 1741 AGACGAGTGAATCTTACGAATCTGAGCTGAGAGACTCTCGGCTTGCTGTAAGAAATTC 1800
QY 1798 AAGCGGAAAGCGAGAAATGTCAGCATTAACGTTGAAGCTTAAGATCAAGGGAAGCTT 1857
Db 1801 AAGCGGAAAGCGAGAAATGTCAGCATTAACGTTGAAGCTTAAGATCAAGGGAAGCTT 1860
QY 1858 GAAAGTGGAGAAATATGCGAAACTGAGAAAGATCAATGCTGAGCAGAGCTCAAAATTCAG 1917
Db 1861 GAAAGTGGAGAAATATGCGAAACTGAGAAAGATCAATGCTGAGCAGAGCTCAAAATTCAG 1920
QY 1918 GAGCTTCAAAGAGAACTGGAAGAGCTGTAAAGCCAGCAAGAGGCTCAGAGCTGCTG 1977
Db 1921 GAGCTTCAAAGAGAACTGGAAGAGCTGTAAAGCCAGCAAGAGGCTCAGAGCTGCTG 1980
QY 1978 CAGAAATATCCGCAAGCAAGAGAGCGAGCCGAGAGAGCTGGAAGAGCTGCAAGAACGGA 2037
Db 1981 CAGAAATATCCGCAAGCAAGAGAGCGAGCCGAGAGAGCTGGAAGAGCTGCAAGAACGGA 2040
QY 2038 GAGGATTCCTTGAAGGCAATCAGAAAGAGCTGTGGAAGCTGAGAAAGCGCCCATTCCT 2097
Db 2041 GAGGATTCCTTGAAGGCAATCAGAAAGAGCTGTGGAAGCTGAGAAAGCGCCCATTCCT 2100
QY 2098 CTGAGAAACAAGTAAAGAGACTAGAGACATGAGCGGTGAAGAAACAAGCTGAAGAT 2157
Db 2101 CTGAGAAACAAGTAAAGAGACTAGAGACATGAGCGGTGAAGAAACAAGCTGAAGAT 2160
QY 2158 GACATCCAGACAAATTCACCAACAGATCCAGCAGATGCTGATTAATTTCTGAGCTCGAA 2217

Dh 2161 GACATCCAGCAAAATCCCAAGATCCAGACATGCTGATTAATTTCTGAGCTGAA 2220
Qy 2218 GAGAAACATCGGAGGAGCCCAAGTCTCAGCCAGCACTAGAGTGCACCTGAAAACAGAAA 2277
Dh 2221 GAGAAACATCGGAGGAGCCCAAGTCTCAGCCAGCACTAGAGTGCACCTGAAAACAGAAA 2280
Qy 2278 GAGCAGCATATGAGAAAAGATTAAAGTTGGACAATCAAGATAAAGAAAGACTGGCT 2337
Dh 2281 GACGACGACATATGAGAAAAGATTAAAGTTGGACAATCAAGATAAAGAAAGACTGGCT 2340
Qy 2338 GACAAGGAGACATGAGGAGACATGATGACAGACAGAGGAGGAGCCCAATGAGAAAGGAG 2397
Dh 2341 GACAAGGAGACATGAGGAGACATGATGACAGACAGAGGAGGAGCCCAATGAGAAAGGAG 2400
Qy 2398 AAAATTCACGCGAACAAGAGCGCATGATCATATGATTCAGATTCAGATCAAGATCCCTG 2457
Dh 2401 AAAATTCACGCGAACAAGAGCGCATGATCATATGATTCAGATTCAGATCAAGATCCCTG 2460
Qy 2458 GAAACAGAGATTGTGAACTGTCTGAAAGCCATTAACCTTGACGAATAGCAGTCTTTTT 2517
Dh 2461 GAAACAGAGATTGTGAACTGTCTGAAAGCCATTAACCTTGACGAATAGCAGTCTTTTT 2520
Qy 2518 ACCCAAGAGACATGAGAGGCGCAAGAGATGATTTCTGAACTCAGGCAACGAAATTT 2577
Dh 2521 ACCCAAGAGACATGAGAGGCGCAAGAGATGATTTCTGAACTCAGGCAACGAAATTT 2580
Qy 2578 TACCTGAGACACAGGCTGGGAAATTGAGAGCCAGAACGAAAACCTGAGAGACAGCTG 2637
Dh 2581 TACCTGAGACACAGGCTGGGAAATTGAGAGCCAGAACGAAAACCTGAGAGAGACAGCTG 2640
Qy 2638 GAGAAATCAGCCACCAAGACCAACAGTACCAAGATCGGCTGCTGAACTGAGACAAAG 2697
Dh 2641 GAGAAATCAGCCACCAAGACCAACAGTACCAAGATCGGCTGCTGAACTGAGACAAAG 2700
Qy 2698 TTGCGGAGAGTCACTCTAGAGACAGAGAGACAGAACTGAGAGCTCAAGCCGCACTCA 2757
Dh 2701 TTGCGGAGAGTCACTCTAGAGACAGAGAGACAGAACTGAGAGCTCAAGCCGCACTCA 2760
Qy 2758 GAGCTACAGCTCTCCCTGACAGAGAGCGGAGTCAAGATTGACAGCCCTGACAGGCTGAC 2817
Dh 2761 GAGCTACAGCTCTCCCTGACAGAGAGCGGAGTCAAGATTGACAGCCCTGACAGGCTGAC 2820
Qy 2818 GCGGCCCCTGAGAGCCAGCTTCGCGCAGCGAGACAGAGCTGAGAGAGACCAACAGCAGAA 2877
Dh 2821 GCGGCCCCTGAGAGCCAGCTTCGCGCAGCGAGACAGAGCTGAGAGAGACCAACAGCAGAA 2880
Qy 2878 GCTGAAAGAGAGATCCAGGCACTCAAGGCACTAGAGATGAATCCAGCCGAATTTGAT 2937
Dh 2881 GCTGAAAGAGAGATCCAGGCACTCAAGGCACTAGAGATGAATCCAGCCGAATTTGAT 2940
Qy 2938 GCTCTTGTAACAGCTGTACTGTAAATCAAGACCTGAGAGAGCGCTAAACAGCTGAC 2997
Dh 2941 GCTCTTGTAACAGCTGTACTGTAAATCAAGACCTGAGAGAGCGCTAAACAGCTGAC 3000
Qy 2998 GAGACACACGCTGAATCTCAACACCAAACTTCTTATCTGTCACAAACAACTCGATGAGCT 3057
Dh 3001 GAGACACACGCTGAATCTCAACACCAAACTTCTTATCTGTCACAAACAACTCGATGAGCT 3060
Qy 3058 TCTGCGCCCAACGACGAGATTGTACAACTGCGAAGTGAAGTGAACATCTCCGCGGAG 3117
Dh 3061 TCTGCGCCCAACGACGAGATTGTGTACAACTGCGAAGTGAAGTGAACATCTCCGCGGAG 3120
Qy 3118 ATACAGGAGAGAGATGAGCTTACAGACAGAGAGAAACGATGAGAGGCTGAGAAAGC 3177
Dh 3121 ATACAGGAGAGAGATGAGCTTACAGACAGAGAGAAACGATGAGAGGCTGAGAAAGC 3180
Qy 3178 ACCTGACACATGCTGAGAGAAACAGATCATGATTTGAGAGGCTTAAACGATGAGCTGTA 3237
Dh 3181 ACCTGACACATGCTGAGAGAAACAGATCATGATTTGAGAGGCTTAAACGATGAGCTGTA 3240
Qy 3238 GAAAAAGAGCGGACGATGAGAGGCTTGAAGAGCGTCTTGAGTGAAGAAATCCAGTTT 3297
Dh 3241 GAAAAAGAGCGGACGATGAGAGGCTTGAAGAGCGTCTTGAGTGAAGAAATCCAGTTT 3300

Qy 3298 GAGTGTGGGTTTGAGAGCTGACAGAGATGCTGACACCGAGAAACAGAGACGGGCGAGA 3357
Dh 3301 GAGTGTGGGTTTGAGAGCTGACAGAGATGCTGACACCGAGAAACAGAGACGGGCGAGA 3360
Qy 3358 GCGGATCAGGAGATCAACCGAGTCTCCGACAGTGTGAGGCTGACAGAGAGACAGCAAG 3417
Dh 3361 GCGGATCAGGAGATCAACCGAGTCTCCGACAGTGTGAGGCTGACAGAGAGAGACAGCAAG 3420
Qy 3418 GCTGAGATTCTCGCTCTGACAGAGGCTCTCAAGAGACAGAGATGAAAGGCCGAGAGCTTC 3477
Dh 3421 GCTGAGATTCTCGCTCTGACAGAGGCTCTCAAGAGACAGAGATGAAAGGCCGAGAGCTTC 3480
Qy 3478 TCTGACAGCTCAATGACCTGAGAGAAAGACATGCTATGCTTGAATGAATGCCGAGAC 3537
Dh 3481 TCTGACAGCTCAATGACCTGAGAGAAAGACATGCTATGCTTGAATGAATGCCGAGAC 3540
Qy 3538 TTAACAGAGAAAGCTGAGAGACATGAAACGAGAGCTCAACAGAGGCTTCTGAGAGAGCAG 3597
Dh 3541 TTAACAGAGAAAGCTGAGAGACATGAAACGAGAGCTCAACAGAGGCTTCTGAGAGAGCAG 3600
Qy 3598 AAATTAACAGAGAGATGAGACCTGACAGAAATCAACATTTCCGCTGATCAAGAGACTG 3657
Dh 3601 AAATTAACAGAGAGATGAGACCTGACAGAAATCAACATTTCCGCTGATCAAGAGACTG 3660
Qy 3658 CAAGAACTCTAGATCGGAGCTGATCTTACTGAGACAGAGAAAGATGAACCTTGAAGTATCAG 3717
Dh 3661 CAAGAACTCTAGATCGGAGCTGATCTTACTGAGACAGAGAAAGATGAAGTACTTGAAGTATCAG 3720
Qy 3718 CTGAGAAACATTCAGGTTCTCTATTTCTCATGAGAAAGGTGAATAATGAGAGCACTATTTCT 3777
Dh 3721 CTGAGAAACATTCAGGTTCTCTATTTCTCATGAGAAAGGTGAATAATGAGAGCACTATTTCT 3780
Qy 3778 CAACAAACCAACTCATGATTTTCTGCAAGCCCAAAATGAGACCAACCTGCTAAAGAAAGAA 3837
Dh 3781 CAACAAACCAACTCATGATTTTCTGCAAGCCCAAAATGAGACCAACCTGCTAAAGAAAGAA 3840
Qy 3838 AAGTTCTCTGACATGACAAATGAGCTGAGCTGAGAGAGAGAGAAAGCTCGCTGT 3897
Dh 3841 AAGTTCTCTGACATGACAAATGAGCTGAGAGCTGAGAGAGAGAGAAAGCTCGCTGT 3900
Qy 3898 GCAAGCTTAGAGAGAGCCCTTCAAGAGACCCGCACTGAGCTCGGCTCCGCCGAGAGAA 3957
Dh 3901 GCAAGCTTAGAGAGAGCCCTTCAAGAGACCCGCACTGAGCTCGGCTCCGCCGAGAGAA 3960
Qy 3958 GCTGCCACCGCAAGAGACAGACCAACCCATCCAGCGCAGCAGCCAGCGAGAG 4017
Dh 3961 GCTGCCACCGCAAGAGACAGACCAACCCATCCAGCGCAGCAGCCAGCGAGAG 4020
Qy 4018 CAGATGCGCATGTCCGCACTGTCGAGTCCGACAGAGACCAAGCCAGTGCATGAGCTG 4077
Dh 4021 CAGATGCGCATGTCCGCACTGTCGAGTCCGACAGAGACCAAGCCAGTGCATGAGCTG 4080
Qy 4078 CTGAGCCCGCCATCCAGCCGAGAGAGAGTCTTCAACCTCAGAGAGATTTAGTGGCGCT 4137
Dh 4081 CTGAGCCCGCCATCCAGCCGAGAGAGAGTCTTCAACCTCAGAGAGATTTAGTGGCGCT 4140
Qy 4138 CTTAAGGAAAGCATGACCAACAATATTCTCACGSAATTCAGAGTGAAGTGAACATGCGA 4197
Dh 4141 CTTAAGGAAAGCATGACCAACAATATTCTCACGSAATTCAGAGTGAAGTGAACATGCGA 4200
Qy 4198 GCGCAAAAGTGTCTGTGTCTGAGATACGATGCACTTTGAGAGCCAGGATCCAAATGT 4257
Dh 4201 GCGCAAAAGTGTCTGTGTCTGAGATACGATGCACTTTGAGAGCCAGGATCCAAATGT 4260
Qy 4258 CTGAATGTCAAGTATGTGTACACCCAGATGCTCAAGTGTCTTGCAGCCACCTGCGGC 4317
Dh 4261 CTGAATGTCAAGTATGTGTACACCCAGATGCTCAAGTGTCTTGCAGCCACCTGCGGC 4320
Qy 4318 TTGCTGCTGAATATGCAACAATTCAACGAGGCTTCTGCGTGAACAAATGAATCTCC 4377
Dh 4321 TTGCTGCTGAATATGCAACAATTCAACGAGGCTTCTGCGTGAACAAATGAATCTCC 4380

4378 CCAGGTCTCCAGACCAAGAGCCGAGAGCCTTGGACCTGAAAGGCTGGATGAAGGTG 4437
4381 CAGGTCTCCAGACCAAGAGCCGAGAGCCTTGGACCTGAAAGGCTGGATGAAGGTG 4440
4438 CCCAGGAATTAACAACGAGAGCAGCAAGGCTGGAGCAAGAAATGATCTGCTGGAGGA 4497
4441 CCCAGGAATTAACAACGAGAGCAGCAAGGCTGGAGCAAGAAATGATCTGCTGGAGGA 4500
4498 TCAGAAATCTCTATTTATGACATGAAGCCAGAGAGCTGGAGCAAGAGCCGCTGGAGAA 4557
4501 TCAGAAATCTCTATTTATGACATGAAGCCAGAGAGCTGGAGCAAGAGCCGCTGGAGAA 4560
4558 TTGAGCTGTGCTTCCGAGCGGGAGATGATATGATGATGCTGCTGGAGTCCGAA 4617
4561 TTGAGCTGTGCTTCCGAGCGGGAGATGATATGATGATGCTGCTGGAGTCCGAA 4620
4618 CTCGCAATTAACAGCAAGAGAGATGCTCCATATGATGAGATGATGATGATGATGATGAT 4677
4621 CTCGCAATTAACAGCAAGAGAGATGCTCCATATGATGAGATGATGATGATGATGATGAT 4680
4678 ACCAGCTGCTGCTCCGAGAGAGATGCTCTATCTGCTGATGCTCCAGCTTCCCTGACAA 4737
4681 ACCAGCTGCTGCTCCGAGAGAGATGCTCTATCTGCTGATGCTCCAGCTTCCCTGACAA 4740
4738 CGCTGGGTCACCGCTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4797
4741 CGCTGGGTCACCGCTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4800
4798 GAAGCTGATGCTTAACTGCTTGGAACTCCCTGCTGAACTGGAGAGTGAAGTGAAGTGA 4857
4801 GAAGCTGATGCTTAACTGCTTGGAACTCCCTGCTGAACTGGAGAGTGAAGTGAAGTGA 4860
4858 GACATGAATCTGACCGCTGCTTCACTGATGACAGAGTGTGTTGGTGGAGAGAGAGAG 4917
4861 GACATGAATCTGACCGCTGCTTCACTGATGACAGAGTGTGTTGGTGGAGAGAGAGAG 4920
4918 CTCTAGCGCTGATGCTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4977
4921 CTCTAGCGCTGATGCTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4980
4978 TTCCAAATTTATTTATTTATCAAGAGCTGAGAGAGCTCTCATGATGATGATGATGATGAT 5037
4981 TTCCAAATTTATTTATTTATTTATCAAGAGCTGAGAGAGCTCTCATGATGATGATGATGAT 5040
5038 GCACTGTGTGTTGGAGCTGAAAGAGTGAAGCAAGTCCCTGGCCCAAGTCCCACTGCTT 5097
5041 GCACTGTGTGTTGGAGCTGAAAGAGTGAAGCAAGTCCCTGGCCCAAGTCCCACTGCTT 5100
5098 GCCCAGCCGAGATCTCAACCCCAATTTTGAAGCTGTCAGAGGCTGCCAATTTGTTGGG 5157
5101 GCCCAGCCGAGATCTCAACCCCAATTTTGAAGCTGTCAGAGGCTGCCAATTTGTTGGG 5160
5158 GCAGGCAAGATTGAGAAAGGAGCTCTGATCTGTCAGAGGCTGCCAAGAGTCCGATTT 5217
5161 GCAGGCAAGATTGAGAAAGGAGCTCTGATCTGTCAGAGGCTGCCAAGAGTCCGATTT 5220
5218 CTCCTGCTACAGAAACCTCAAGCAATATCTGATCTGGAAAGATAGAGACTTCAGAG 5277
5221 CTCCTGCTACAGAAACCTCAAGCAATATCTGATCTGGAAAGATAGAGACTTCAGAG 5280
5278 CCTGAGAGTATCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACTT 5337
5281 CCTGAGAGTATCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACTT 5340
5338 GAATATGACATGAAGCAGTACAGCTGAGAGATTCCTGGATGAAGATGATGATGATGATGAT 5397
5341 GAATATGACATGAAGCAGTACAGCTGAGAGATTCCTGGATGAAGATGATGATGATGATGAT 5400
5398 GCACCTGCTGTGTTGGCTGCTTTCACACAGCTTCCCTGCTCATGCTGTCAGAGTGAAC 5457
5401 GCACCTGCTGTGTTGGCTGCTTTCACACAGCTTCCCTGCTCATGCTGTCAGAGTGAAC 5460
5458 AGCGAGAGGAGAGAGAGATCTTGTGCTTTCACAGATTTGGAGTGTTCGTGAT 5517

5461 AGCGAGGAGAGAGAGAGATCTTGTGTTTCCACCAATTTGAGTGTGATGAT 5520
5518 TCTTACGAGAGAGAGAGAGAGATCTTGTGTTTCCACCAATTTGAGTGTGATGAT 5577
5521 TCTTACGAGAGAGAGAGAGAGATCTTGTGTTTCCACCAATTTGAGTGTGATGAT 5580
5578 GCTTACGAGAGAGAGAGAGATCTTGTGTTTCCACCAATTTGAGTGTGATGATGAT 5637
5581 GCTTACGAGAGAGAGAGAGATCTTGTGTTTCCACCAATTTGAGTGTGATGATGAT 5640
5638 CAGGAG 5697
5641 CAGGAG 5700
5698 TACCTGAG 5757
5701 TACCTGAG 5760
5758 TTAAGGATATTTGCTGCAAG 5817
5761 TTAAGGATATTTGCTGCAAG 5820
5818 GGCCTGTCACCTCCGAG 5877
5821 GGCCTGTCACCTCCGAG 5880
5878 ATACCAAG 5937
5881 ATACCAAG 5940
5941 CCAAGCAG 6000
5998 GGCCTGTCACCTCCGAG 6057
6001 GGCCTGTCACCTCCGAG 6060
6058 TCCCTGAG 6117
6061 TCCCTGAG 6120
6118 ACCCGCTGTCTCCAGGTGAACAAGTCTGGAGCAAGTTC 6155
6121 ACCCGCTGTCTCCAGGTGAACAAGTCTGGAGCAAGTTC 6158
RESULT 8
AX503780 6189 bp DNA linear PAT 27-SEP-2002
LOCUS AX503780
DEFINITION Sequence 10 from Patent W00226826.
ACCESSION AX503780
VERSION AX503780.1 GI:23385965
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE
1 Gerlach, V.L., Macdougall, J.R., Smithson, G., Millet, I., Stone, D.,
Gunther, E., Eilerman, K., Grose, W.M., Alsbrook, J.P., Lepley, D.M.,
Burgess, C.E., Padigaru, M., Kekuda, R., Spytek, K.A., Leach, M.D. and
Shimkets, R.A.
Proteins and nucleic acids encoding same
Patent: WO 0226826-A 10 04-APR-2002;
Curagen Corporation (US)
FEATURES
source
1. 6189
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN

Query Match 99.3%; Score 6121.8; DB 6; Length 6189;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 6145; Conservative 0; Mismatches 17; Indels 3; Gaps 1;

QY 1 ATGTGAGTTCAAATATGAGGCGGAATCTTTGGATGCTGTGCTGAACCCATT 60
DB 1 ATGTGAGTTCAAATATGAGGCGGAATCTTTGGATGCTGTGCTGAACCCATT 60
QY 61 GCCAGCGGCGCTCCAGGCTGAATCTGTCTCCAGGGGAAACCAACCTTATGACTCA 120
DB 61 GCCAGCGGCGCTCCAGGCTGAATCTGTCTCCAGGGGAAACCAACCTTATGACTCA 120
QY 121 CACAGATGTCTCTCTTTCCGAGAAAGGATATTAGATCCCTCTTTGTCTTTGAA 180
DB 121 CACAGATGTCTCTCTTTCCGAGAAAGGATATTAGATCCCTCTTTGTCTTTGAA 180
QY 181 GAATGCACTGAGCTGTCTGATGAGATTTAGACCTGAGCACTTTGCCGAAGTAT 240
DB 181 GAATGCACTGAGCTGTCTGATGAGATTTAGACCTGAGCACTTTGCCGAAGTAT 240
QY 241 TCCGACACCATAGCTGATTAACAGAGCTCAGGCTTCGGCAAGACTTCGAAGTACA 300
DB 241 TCCGACACCATAGCTGATTAACAGAGCTCAGGCTTCGGCAAGACTTCGAAGTACA 300
QY 301 AGCTTGTAGTGTGTGTCACTTTGTGAAGTGAAGTGTGTGAAGAGAAACCGGG 360
DB 301 AGCTTGTAGTGTGTGTGTCACTTTGTGAAGTGAAGTGTGTGAAGAGAAACCGGG 360
QY 361 GACATCTATGCTATGAAAGTGAAGAGAAAGAGCTTTATTTGGCCAGAGCTTTCA 420
DB 361 GACATCTATGCTATGAAAGTGAAGAGAAAGAGCTTTATTTGGCCAGAGCTTTCA 420
QY 421 TTTTGTGAGGAAGGCGGAACATTTATCTGAAGCAAGCCCGTGAATCCCAATTA 480
DB 421 TTTTGTGAGGAAGGCGGAACATTTATCTGAAGCAAGCCCGTGAATCCCAATTA 480
QY 481 CAGATGCTCTTTCAGGACAAATACCTTTATCTGTGATGAGATATCAGCTGAGGG 540
DB 481 CAGATGCTCTTTCAGGACAAATACCTTTATCTGTGATGAGATATCAGCTGAGGG 540
QY 541 GACTGTGCTGACTTTGAAATAGATGAGGACAGTTAGTGAATGAACTGATACGTTT 600
DB 541 GACTGTGCTGACTTTGAAATAGATGAGGACAGTTAGTGAATGAACTGATACGTTT 600
QY 601 TACTGAGTGAAGTATTTGGCTGTCAAGGCTTATCGATGAGGATACGTGCATCGA 660
DB 601 TACTGAGTGAAGTATTTGGCTGTGTCAAGGCTTATCGATGAGGATACGTGCATCG 660
QY 661 GACATCAAGCTGAGAACTTTCTGTTGACGACAGAGACATCAAGCTGATGATTTT 720
DB 661 GACATCAAGCTGAGAACTTTCTGTTGACGACAGAGACATCAAGCTGATGATTTT 720
QY 721 GGAATCTGCGCGAAATGAAATTCACAGATGATGATCCCAATCTCCGATTTGGACC 780
DB 721 GGAATCTGCGCGAAATGAAATTCACAA---GGTGAATGCAAACTCCGATTTGGACC 777
QY 781 CCAATTAACATGCTCTCTGAAGTCTGATGATGAACGGGATGGAAGAGCACTTAC 840
DB 781 CCAATTAACATGCTCTCTGAAGTCTGATGATGAACGGGATGGAAGAGCACTTAC 837
QY 841 GGCCTGAGCTGTGACTGTGATGATGATGATGATGATGATGATGATGATGATGATG 900
DB 841 GGCCTGAGCTGTGACTGTGATGATGATGATGATGATGATGATGATGATGATGATG 897
QY 901 TCCCTCTTCCGAGGGAACCTTGCAGAACTTCAATACATTAATGAATTTCCAGCGG 960
DB 901 TCCCTCTTCCGAGGGAACCTTGCAGAACTTCAATACATTAATGAATTTCCAGCGG 957
QY 961 TTTTGAATTTTCCAGATGAGCCCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1020
DB 961 TTTTGAATTTTCCAGATGAGCCCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1017

QY 1021 TTGTGTGCGGCGAGAAAGAGAGACTGAAGTTTGAAGTCTTTCGTCATCTTCTTC 1080
DB 1018 TTGTGTGTGCGGCGAGAAAGAGAGACTGAAGTTTGAAGTCTTTCGTCATCTTCTTC 1077
QY 1081 TCTAAATTTGACTGGAACAACTTTCGTAATCTCTCTCCCTTCTGTTCCCACTTCAAG 1140
DB 1078 TCTAAATTTGACTGGAACAACTTTCGTAATCTCTCTCCCTTCTGTTCCCACTTCAAG 1137
QY 1141 TCTGAGATGACACTTCATTTTGAATGAACAGAGAAAGATTCGTGGTTCATCTCT 1200
DB 1138 TCTGAGATGACACTTCATTTTGAATGAACAGAGAAAGATTCGTGGTTCATCTCT 1197
QY 1201 CCGTGCAGCTGAGCCCTTCAAGCTTCTCGGGTGAAGAACTGCCCTTGTGGGGTTTTCG 1260
DB 1198 CCGTGCAGCTGAGCCCTTCAAGCTTCTCGGGTGAAGAACTGCCCTTGTGGGGTTTTCG 1257
QY 1261 TACAGAAAGGCACTGGGGATTTCTTGGATGATGAGTCTGTGTGTGCTGAGCTCC 1320
DB 1258 TACAGAAAGGCACTGGGGATTTCTTGGATGATGAGTCTGTGTGTGCTGAGCTCC 1317
QY 1321 CTTGCCAAGACTAGCTTCATGAGAAAGAACTTCTCATCAAAAGCAAGAGCTTCAAGAC 1380
DB 1318 CTTGCCAAGACTAGCTTCATGAGAAAGAACTTCTCATCAAAAGCAAGAGCTTCAAGAC 1377
QY 1381 TCTCAGAGCAAGTGTCAAAAGATGAGAGAGAAATGACCCGTTTACATCGGAAGTCTCA 1440
DB 1378 TCTCAGAGCAAGTGTCAAAAGATGAGAGAGAAATGACCCGTTTACATCGGAAGTCTCA 1437
QY 1441 GAGTGAAGCTGTGCTTGTGTGAGAGAGAGTGAAGGCTCTGAGACTCAGAGA 1500
DB 1438 GAGTGAAGCTGTGCTTGTGTGAGAGAGTGAAGGCTCTGAGACTCAGAGA 1497
QY 1501 TCCCTCTGAGAGAGACCTTGTCTTACTCATCAGATGAGTGAAGCTTTAAAGCAAGT 1560
DB 1498 TCCCTCTGAGAGAGACCTTGTCTTACTCATCAGATGAGTGAAGCTTTAAAGCAAGT 1557
QY 1561 TTGAGACAGCAAGTGTGAGTGTCTCAGAGAGATGACAAAGCTGACGCTTCTCAT 1620
DB 1558 TTGAGACAGCAAGTGTGAGTGTCTCAGAGAGATGACAAAGCTGACGCTTCTCAT 1617
QY 1621 GATATCAGAGAGAGAGCGGAGCTTCAGAAATCAAGAGAGAGTCCAGGCTCAA 1680
DB 1618 GATATCAGAGAGAGAGCGGAGCTTCAGAAATCAAGAGAGAGTCCAGGCTCAA 1677
QY 1681 GTGAGAAATGAGTGTGATGATGAATCAAGTTGAAAGAGATCTTGTCTCAGCAAGAGA 1740
DB 1678 GTGAGAAATGAGTGTGATGATGAATCAAGTTGAAAGAGATCTTGTCTCAGCAAGAGA 1737
QY 1741 CCGAGTATCTTACGAATCTGAGCTGAGAGACTTGGCTTGTCTGTGAAGATTCAG 1800
DB 1738 CCGAGTATCTTACGAATCTGAGCTGAGAGACTTGGCTTGTCTGTGAAGATTCAG 1797
QY 1801 CCGAAAGCGACAGATGTCAAGCTTAACTTTGAAGCTTAAGATCAAGGAAAGCTTGAA 1860
DB 1798 CCGAAAGCGACAGATGTCAAGCTTAACTTTGAAGCTTAAGATCAAGGAAAGCTTGAA 1857
QY 1861 GTGAGAAATGAGTGTGAGAGATCAATGCTGAGAGAGCTCAAAATTCAGAG 1920
DB 1858 GTGAGAAATGAGTGTGAGAGATCAATGCTGAGAGAGCTCAAAATTCAGAG 1917
QY 1921 CTCAGAGAGAACTGAGAGAGCTGTAAAGCCAGCACTGAGGCGCACAGCTGTCTCAG 1980
DB 1918 CTCAGAGAGAACTGAGAGAGCTGTAAAGCCAGCACTGAGGCGCACAGCTGTCTCAG 1977
QY 1981 AATATCCGCAAGGCAAGAGCGAGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2040
DB 1978 AATATCCGCAAGGCAAGAGCGAGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2037
QY 2041 GATTTCTTGAAGGCAATCAGAAAGAGCTGTGAAAGCTGAGGAAAGCCGCCCATTTCTCTG 2100
DB 2038 GATTTCTTGAAGGCAATCAGAAAGAGCTGTGAAAGCTGAGGAAAGCCGCCCATTTCTCTG 2097
QY 2101 GAGAACAAAGGTAAAGAGACTTGAAGACCATGAGCGTGAAGAAACAGACTGAAGATGAC 2160

```
Db      2098 GAGAACTGTAAGGAGCTGAGACCATGAGGCTTGAAGAAACAGCTGAGAGATGAC      2157
Qy      2161 ATCCAGCAAAAATCCCAACAGATCAGCAGATGAGCTGATATAAATTTCTGAGCTCGAAGAG      2220
Db      2158 ATCCAGCAAAAATCCCAACAGATCAGCAGATGAGCTGATATAAATTTCTGAGCTCGAAGAG      2217
Qy      2221 AAACATGGGAGGGCCCAAGTCTCAGCCCAAGACCTTAGAAGTGCACCTGAAAAGAGAG      2280
Db      2218 AAACATGGGAGGGCCCAAGTCTCAGCCCAAGACCTTAGAAGTGCACCTGAAAAGAGAG      2277
Qy      2281 CAGCACTATGAGGAAAAGATTAAAGTGTGGAACAATCAGATTAAGAAAGACCTGGCTGAC      2340
Db      2278 CAGCACTATGAGGAAAAGATTAAAGTGTGGAACAATCAGATTAAGAAAGACCTGGCTGAC      2337
Qy      2341 AAGGAGCACTGAGGAACATGATGACAGACACGAGGAGAGAGGCCATGAAAGGCGAAA      2400
Db      2338 AAGGAGCACTGAGGAACATGATGACAGACACGAGGAGAGAGGCCATGAAAGGCGAAA      2397
Qy      2401 ATTCTCAGCGGAACGAAAGGCGATGATCAATGCTATGGAATTCGAAGTCCAGATCCCTGAAA      2460
Db      2398 ATTCTCAGCGGAACGAAAGGCGATGATCAATGCTATGGAATTCGAAGTCCAGATCCCTGAAA      2457
Qy      2461 CAGAGATTTGGAATCTGTCTGAAGCCATTAACCTTGACGAAATAGCAGTCTTTTAC      2520
Db      2458 CAGAGATTTGGAATCTGTCTGAAGCCATTAACCTTGACGAAATAGCAGTCTTTTAC      2517
Qy      2521 CAAGGAAACATGAAGGCCCAAGAAAGATGATTTCTGAACTCAGGCAACGAAATTTTAC      2580
Db      2518 CAAGGAAACATGAAGGCCCAAGAAAGATGATTTCTGAACTCAGGCAACGAAATTTTAC      2577
Qy      2581 CTGGGAACACAGGCTGGGAAGTTGGAGGCCCAAGAACGGAACCTGGAGAGACGCTGGAG      2640
Db      2578 CTGGGAACACAGGCTGGGAAGTTGGAGGCCCAAGAACGGAACCTGGAGAGACGCTGGAG      2637
Qy      2641 AAGATCAGGCCCAAGAAACAAGTGAACAAGATCGGCTGGAACCTGGAGAGCAAGATTG      2700
Db      2638 AAGATCAGGCCCAAGAAACAAGTGAACAAGATCGGCTGGAACCTGGAGAGCAAGATTG      2697
Qy      2701 CGGAGAGTCACTGTAGAGACGAGAGACAGAACTGGAGCTCAAGGCCACAGTCAAGAG      2760
Db      2698 CGGAGAGTCACTGTAGAGACGAGAGACAGAACTGGAGCTCAAGGCCACAGTCAAGAG      2757
Qy      2761 CTACAGCTCTCCCTGAGAGGCCGAGTCAAGTGAACAGCTCTGACAGCTGCAAGCGCG      2820
Db      2758 CTACAGCTCTCCCTGAGAGGCCGAGTCAAGTGAACAGCTCTGACAGCTGCAAGCGCG      2817
Qy      2821 GCCCTGAGAGCCAGGCTGCGCAGGGGAAAGACAGAGCTGGAAGAGACACAGAGAAAGCT      2880
Db      2818 GCCCTGAGAGCCAGGCTGCGCAGGGGAAAGACAGAGCTGGAAGAGACACAGAGAAAGCT      2877
Qy      2881 GAAAGAGAGATCCAGGCACTCAGGCAATAGAGATGAATCCAGGCGAAATTTGATGCT      2940
Db      2878 GAAAGAGAGATCCAGGCACTCAGGCAATAGAGATGAATCCAGGCGAAATTTGATGCT      2937
Qy      2941 CTTTCGTAACAGCTGTACTGTATCAAGACCTGGAGAGAGAGCTTAAACCAAGCTGACCGAG      3000
Db      2938 CTTTCGTAACAGCTGTACTGTATCAAGACCTGGAGAGAGAGCTTAAACCAAGCTGACCGAG      2997
Qy      3001 GACAAAGCTGAACCTCAACCAACCAAACTTCTACTGTCCAAACAATCGATAGAGGCTTCT      3060
Db      2998 GACAAAGCTGAACCTCAACCAACCAAACTTCTACTGTCCAAACAATCGATAGAGGCTTCT      3057
Qy      3061 GCGCGCAACGAGAGATTTGTACACTGCGAAGTGAAGTGAACCATCTCCGCGGAGATC      3120
Db      3058 GCGCGCAACGAGAGATTTGTACACTGCGAAGTGAAGTGAACCATCTCCGCGGAGATC      3117
Qy      3121 ACGGAACGAGAGATGAGCTTACCAAGCCAGCAAGAAACGATGAGGCTGTGAAGACACG      3180
Db      3118 ACGGAACGAGAGATGAGCTTACCAAGCCAGCAAGAAACGATGAGGCTGTGAAGACACG      3177
Qy      3181 TGCAACATGCTGAGAGACAGTCAATGATTTGGAGGCCCTTAAACGATGAGCTGTAGAA      3240
Db      3178 TGCAACATGCTGAGAGACAGTCAATGATTTGGAGGCCCTTAAACGATGAGCTGTAGAA      3237
Qy      3241 AAAGCGGAGAGTGGAGGCGCTGAGAGAGCGCTCTGGGTGATGAGAAATCCAGTTTGA      3300
Db      3238 AAAGCGGAGAGTGGAGGCGCTGAGAGAGCGCTCTGGGTGATGAGAAATCCAGTTTGA      3297
Qy      3301 TGTGCGGTTTGAAGACTGCGAGAGATGCTGACACCGAGAAACAGACAGGCGGAGAGCC      3360
Db      3298 TGTGCGGTTTGAAGACTGCGAGAGATGCTGACACCGAGAAACAGACAGGCGGAGAGCC      3357
Qy      3361 GATCAGCGATCAACGAGCTCGCAGGAGTGGAGAGCTGACAGTGAAGAGACCAAGGCT      3420
Db      3358 GATCAGCGATCAACGAGCTCGCAGGAGTGGAGAGCTGACAGTGAAGAGACCAAGGCT      3417
Qy      3421 GAGATTTCTGCTCTGAGCAGAGCTCTCAAGAGCAGAAAGCTGAAGGCCAGAGCTCTCT      3480
Db      3418 GAGATTTCTGCTCTGAGCAGAGCTCTCAAGAGCAGAAAGCTGAAGGCCAGAGCTCTCT      3477
Qy      3481 GACAGCTCAATGACTCTGAGAGAAAGCATGCTATGCTGAAATGAATGCCAAGCTTA      3540
Db      3478 GACAGCTCAATGACTCTGAGAGAAAGCATGCTATGCTGAAATGAATGCCAAGCTTA      3537
Qy      3541 CAGCAGAGCTGAGACCTGAACGAGAGCTCAACAGAGGCTTCTGGAAGAGCAAGCCAA      3600
Db      3538 CAGCAGAGCTGAGAGCTGAACGAGAGCTCAACAGAGGCTTCTGGAAGAGCAAGCCAA      3597
Qy      3601 TTACAGCAGCAGATGAGACCTGCAGAAATAATCAATTTTCCTGTGACTCAAGAGCTGCA      3660
Db      3598 TTACAGCAGCAGATGAGACCTGCAGAAATAATCAATTTTCCTGTGACTCAAGAGCTGCA      3657
Qy      3661 GAAGCTTGAATGCGGCTGATCTAAGAGACAGAAAGAGATGATCTTGAAGATACAGCTG      3720
Db      3658 GAAGCTTGAATGCGGCTGATCTAAGAGACAGAAAGAGATGATCTTGAAGATACAGCTG      3717
Qy      3721 GAAAAATTCAGGTTCTCAATCTCAATGAAAGGTGAAATGGAAGGCACTAATTTCTCA      3780
Db      3718 GAAAAATTCAGGTTCTCAATCTCAATGAAAGGTGAAATGGAAGGCACTAATTTCTCA      3777
Qy      3781 CAAACCAATCTCAATTTTCTGCAAGGCCAAATGAGCAACCTCTTAAAAAGAAAAAG      3840
Db      3778 CAAACCAATCTCAATTTTCTGCAAGGCCAAATGAGCAACCTCTTAAAAAGAAAAAG      3837
Qy      3841 GTTCTCTGCAATCAATGAGCTGAAAGCTGGCCCTGGAGAGAGAAAGCTGCTGTGCA      3900
Db      3838 GTTCTCTGCAATCAATGAGCTGAAAGCTGGCCCTGGAGAGAGAAAGCTGCTGTGCA      3897
Qy      3901 GAGCTTGAAGAGCCCTTGAAGAACCCGCAATCGAGCTCCGCTCCGCGGAGAGAGCT      3960
Db      3898 GAGCTTGAAGAGCCCTTGAAGAACCCGCAATCGAGCTCCGCTCCGCGGAGAGAGCT      3957
Qy      3961 GCCACCGGAAAGCAACGAGACCAACCAATCCATCCAGCCAGCCACCGCGAGAGAGAG      4020
Db      3958 GCCACCGGAAAGCAACGAGACCAACCAATCCATCCAGCCAGCCACCGCGAGAGAGAG      4017
Qy      4021 ATGCGCATGTCCGCAATCGTGGGTGCGAGACAGACACAGCCCAAGTGCATGAGCTGTG      4080
Db      4018 ATGCGCATGTCCGCAATCGTGGGTGCGAGACAGACACAGCCCAAGTGCATGAGCTGTG      4077
Qy      4081 GCGCGGCAATCCAGCGCGAGAAAGAGTCTTAACTCCAGAGAAATTTAGTGGGCTCTT      4140
Db      4078 GCGCGGCAATCCAGCGCGAGAAAGAGTCTTAACTCCAGAGAAATTTAGTGGGCTCTT      4137
Qy      4141 AAGGAACGATGACCAACATTTCTCAACCGATTCAGCTGAGACTGAACATGCGAGCC      4200
Db      4138 AAGGAACGATGACCAACATTTCTCAACCGATTCAGCTGAGACTGAACATGCGAGCC      4197
Qy      4201 ACAAAAGTGTGTGTGTGTGTGATACCGTGAACCTTTGAACGCGCAGAGATCCAAATGTCTC      4260
Db      4198 ACAAAAGTGTGTGTGTGTGTGATACCGTGAACCTTTGAACGCGCAGAGATCCAAATGTCTC      4257
Qy      4261 GAATGTCAAGTGAATGTGACCCCAAGTGTCCAGGTCTTGGCAGCCACTGGCGGCTTG      4320
Db      4258 GAATGTCAAGTGAATGTGACCCCAAGTGTCCAGGTCTTGGCAGCCACTGGCGGCTTG      4317
```

QY 4321 CTTGCTGAATATGCGACACACTTCAACGAGGCTTTCTGCGGTGCAAAATGAATCTCCCA 4380
 DB 4318 CTTGCTGAATATGCGACACACTTCAACGAGGCTTTCTGCGGTGCAAAATGAATCTCCCA 4377
 QY 4381 GGTCTCCAGACCAAGAGGCCAGACGACTTGTGACCTTGGAAAGGTGTGAATGAGTGTCC 4440
 DB 4378 GGTCTCCAGACCAAGAGGCCAGACGACTTGTGACCTTGGAAAGGTGTGAATGAGTGTCC 4437
 QY 4441 AGGAATTAACAAACGAGGACGAGGCTGGGACAGAAAGTACATTTGCTTGAAGGATCA 4500
 DB 4438 AGGAATTAACAAACGAGGACGAGGCTGGGACAGAAAGTACATTTGCTTGAAGGATCA 4497
 QY 4501 AAAGTCTCATTTATGACAAAGGACGAGGACGAGGCTGGGACAGAAAGTATTT 4560
 DB 4498 AAAGTCTCATTTATGACAAAGGACGAGGACGAGGCTGGGACAGAAAGTATTT 4557
 QY 4561 GAGCTGTGCTTTCCCAACGAGGAGTGTATCTATTTGAGTGTGCTTCCGAACTC 4620
 DB 4558 GAGCTGTGCTTTCCCAACGAGGAGTGTATCTATTTGAGTGTGCTTCCGAACTC 4617
 QY 4621 GCAAAATACAGCAAGAGGATGTCCCATACATCTGAAGATGGAATCTCAACCGCACACC 4680
 DB 4618 GCAAAATACAGCAAGAGGATGTCCCATACATCTGAAGATGGAATCTCAACCGCACACC 4677
 QY 4681 ACCGTGTGACCCGAGGAGACCTTCTACTGTGCTGACCTCCAGCTTCCCTGACAAACAGCGC 4740
 DB 4678 ACCGTGTGACCCGAGGAGACCTTCTACTGTGCTGACCTCCAGCTTCCCTGACAAACAGCGC 4737
 QY 4741 TGGGTACCGCTTTAGATCATGTTGTGCGAGGTGGAGAGTTTCTAGGAAAAAGCGAA 4800
 DB 4738 TGGGTACCGCTTTAGATCATGTTGTGCGAGGTGGAGAGTTTCTAGGAAAAAGCGAA 4797
 QY 4801 GCTGATCTAAACGCTTGGAACTCCCTGCTGAACCTGGAAGGTGAAGA CCGTCTAGAC 4860
 DB 4798 GCTGATCTAAACGCTTGGAACTCCCTGCTGAACCTGGAAGGTGAAGA CCGTCTAGAC 4857
 QY 4861 ATGAATGACAGCTGCTCCCTTCACTGACCAAGTGTGTGTGAGGACCGAGAAAGGCTC 4920
 DB 4858 ATGAATGACAGCTGCTCCCTTCACTGACCAAGTGTGTGTGAGGACCGAGAAAGGCTC 4917
 QY 4921 TACGCCCTGAATGTCTGAAAAAATCCCTTAACCCATGTCCGAGAAATTTGAGACAGTCTTC 4980
 DB 4918 TACGCCCTGAATGTCTGAAAAAATCCCTTAACCCATGTCCGAGAAATTTGAGACAGTCTTC 4977
 QY 4981 CAAATTTATTTATTCAGAGACCTGGAAGTACTCATGATGAGGATGAAGAGAGCGGACA 5040
 DB 4978 CAAATTTATTTATTCAGAGACCTGGAAGTACTCATGATGAGGATGAAGAGAGCGGACA 5037
 QY 5041 CTGTGTCTTGTGACGTAAGAAAGTGAACAGTCCCTGGGCCAGTCCCACTGCTGCC 5100
 DB 5038 CTGTGTCTTGTGACGTAAGAAAGTGAACAGTCCCTGGGCCAGTCCCACTGCTGCC 5097
 QY 5101 CAGCCGACATCTCAACCAACATTTTGAAGCTGTCAAGGCTGCACTTGTGTGGGACA 5160
 DB 5098 CAGCCGACATCTCAACCAACATTTTGAAGCTGTCAAGGCTGCACTTGTGTGGGACA 5157
 QY 5161 GCGAAGATTGAGAACGGGCTGTGATCTGTGACACCATGCCCCAGCAAAAGTGTCAATCTC 5220
 DB 5158 GCGAAGATTGAGAACGGGCTGTGATCTGTGACACCATGCCCCAGCAAAAGTGTCAATCTC 5217
 QY 5221 CGCTACACGAAAACTCAGCAATATCTGACCTCGGAAAGAGATAGAGACCTCAGAGCCC 5280
 DB 5218 CGCTACACGAAAACTCAGCAATATCTGACCTCGGAAAGAGATAGAGACCTCAGAGCCC 5277
 QY 5281 TGCAGCTGTATCACTTACCAATTTACATGATCTCATTTGAAACCAATTAATTTACGAA 5340
 DB 5278 TGCAGCTGTATCACTTACCAATTTACATGATCTCATTTGAAACCAATTAATTTACGAA 5337
 QY 5341 ATGCAATGAAGAGTACAGCTCGAGGAATTTCTGTGATAGAAATGACCAATTTCTTGGCA 5400
 DB 5338 ATGCAATGAAGAGTACAGCTCGAGGAATTTCTGTGATAGAAATGACCAATTTCTTGGCA 5397

QY 5401 CTTGCTGTGTTTGGCCGCTCTTCCAAACAGCTTCCCTGTCTCATCTGCAAGTGAACAGC 5460
 DB 5398 CTTGCTGTGTTTGGCCGCTCTTCCAAACAGCTTCCCTGTCTCATCTGCAAGTGAACAGC 5457
 QY 5461 GCAAGGACGAGAGGAGATCTTGTGTTTCCAGAAATTTGGAGTGTGTGTGATTTCT 5520
 DB 5458 GCAAGGACGAGAGGAGATCTTGTGTTTCCAGAAATTTGGAGTGTGTGTGATTTCT 5517
 QY 5521 TACGAGACGTAAGCCGACAGACGATCTCAAGTGAAGTGTGCTTACCTTTGGCTTTGCC 5580
 DB 5518 TACGAGACGTAAGCCGACAGACGATCTCAAGTGAAGTGTGCTTACCTTTGGCTTTGCC 5577
 QY 5581 TACAGAGAACCTTATCTGTTTGTGACCCACTTCACTCACTCGAAGTAATTGATGATCAG 5640
 DB 5578 TACAGAGAACCTTATCTGTTTGTGACCCACTTCACTCACTCGAAGTAATTGATGATCAG 5637
 QY 5641 GCAAGCTCTCAGAGGAGACCCGCTGCGGACGATCTGAGACATCCGGAACCCGCGCTAC 5700
 DB 5638 GCAAGCTCTCAGAGGAGACCCGCTGCGGACGATCTGAGACATCCGGAACCCGCGCTAC 5697
 QY 5701 CTGGGCTCTGCAATTTCTCAGAGAGCGATTTACTTGTGCTTCTCATACAGATTAATTA 5760
 DB 5698 CTGGGCTCTGCAATTTCTCAGAGAGCGATTTACTTGTGCTTCTCATACAGATTAATTA 5757
 QY 5761 AGGCTCATTTTGTGCAAGGAAACCTCTGTGAAGAGTCCGACACTGAACACACCGGAGC 5820
 DB 5758 AGGCTCATTTTGTGCAAGGAAACCTCTGTGAAGAGTCCGACACTGAACACACCGGAGC 5817
 QY 5821 CCGTCCACCTCCCGACAGACCCCAACAGAGGCGCCACCGTCAACAGACACATC 5880
 DB 5818 CCGTCCACCTCCCGACAGACCCCAACAGAGGCGCCACCGTCAACAGACACATC 5877
 QY 5881 ACCAAGCGCTGACCTTCCAGCCAGCGCCGAGAGGCGCCAGCCAGCCAGGAGGAGCA 5940
 DB 5878 ACCAAGCGCTGACCTTCCAGCCAGCGCCGAGAGGCGCCAGCCAGCCAGGAGGAGCA 5937
 QY 5941 AGCACACCCACCGCTACCCGAGAGGCGGAGCGAGCTGCGAGAGCAAGTCTCTGGC 6000
 DB 5938 AGCACACCCACCGCTACCCGAGAGGCGGAGCGAGCTGCGAGAGCAAGTCTCTGGC 5997
 QY 6001 CGCCCCCTGAGAGGAGAAAGTCCCCCGGCGGATATCTGAGCAACCGGAGAGAGCGGTCC 6060
 DB 5998 CGCCCCCTGAGAGGAGAAAGTCCCCCGGCGGATATCTGAGCAACCGGAGAGAGCGGTCC 6057
 QY 6061 CCGCGAGGCTGTTTGAAGA CAGCAGAGGAGGCGGCTGCTGCGGAGAGCGGTGAGAGC 6120
 DB 6058 CCGCGAGGCTGTTTGAAGA CAGCAGAGGAGGCGGCTGCTGCGGAGAGCGGTGAGAGC 6117
 QY 6121 CCGCTGTCCAGGTGAACAAAGTCTGAGACCAAGTCTTCAATATA 6165
 DB 6118 CCGCTGTCCAGGTGAACAAAGTCTGAGACCAAGTCTTCAATATA 6162

RESULT 9
 AX503778
 LOCUS 6201 bp DNA linear PAT 27-SEP-2002
 DEFINITION Sequence 8 from Patent WO0226826.
 AX503778
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Homo sapiens (human)
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
 AUTHORS
 1 Gerlach,V.L., Macdougall,J.R., Smithson,G., Milliet,I., Stone,D.,
 Gunther,E., Ellerman,K., Grose,W.M., Alsbrook,J.P., Lepley,D.M.,
 Burgess,C.E., Padigar,M., Kékuda,R., Spyek,K.A., Leach,M.D. and
 Shimkets,R.A.
 Proteins and nucleic acids encoding same
 Patent: WO 0226826-A 8 04-APR-2002;
 Curagen Corporation (US)
 FEATURES
 Location/Qualifiers

source	1. .6201	/organism="Homo sapiens"	/mol_type="unassigned DNA"	/db_xref="taxon:9606"
ORIGIN				
Query Match	99.0%	Score 6105.4	DB 6	Length 6201
Best Local Similarity	99.6%	Pred. No. 0		
Matches 6131	Conservative	0	Mismatches 21	Indels 3
			Gaps	1
QY	1	AGTTGAAGTTCAAATATATGAGACGGGGAAATCCTTTGATAGCTGCTGCTGACACCACTT	60	
DB	1	ATGTTGAACTTCAAAATATGAGACGGGGAAATCCTTTGATAGCTGCTGCTGACACCACTT	60	
QY	61	GCCAGCCGGGCTCCAGGCTGAAATCTGTTCTTCCAGGGGAAACCACTTTATGACTCAA	120	
DB	61	GCCAGCCGGGCTCCAGGCTGAAATCTGTTCTTCCAGGGGAAACCACTTTATGACTCAA	120	
QY	121	CAGCAGATGTCCTCTCTTTCCCGAAGAGATATTAATGCCCTCTTGTCTCTTGAA	180	
DB	121	CAGCAGATGTCCTCTCTTTCCCGAAGAGATATTAATGATGCCCTCTTGTCTCTTGAA	180	
QY	181	GAATGACATGCGCCGCTCTGAATGAATTAAGCACTGACGAACTTTGTCGGAAATAT	240	
DB	181	GAATGACATGCGCCGCTCTCTGAATGAATTAAGCACTGACGAACTTTGTCGGAAATAT	240	
QY	241	TCCGACACCATATGACTGAGTTACAGAGCTCCAGCCTTGGGAAAGGAACTTGGAAATCAGA	300	
DB	241	TCCGACACCATATGACTGAGTTACAGAGCTCCAGCCTTGGGAAAGGAACTTGGAAATCAGA	300	
QY	301	AGCTCTTGAAGTTGTGTCTCACTTGTCTGAAGTGCAGTGTGTGAAGAGAAAGCAACGGG	360	
DB	301	AGCTCTTGAAGTTGTGTGTCTCACTTGTCTGAAGTGCAGTGTGTGAAGAGAAAGCAACGGG	360	
QY	361	GACATCTATGCTATGAAGAATGAAGAAGAGCTTTATATGGCCAGAGACAGTTTCA	420	
DB	361	GACATCTATGCTATGAAGAATGAAGAAGAGCTTTATATGGCCAGAGACAGTTTCA	420	
QY	421	TTTTTGAAGAAAGCGGACATATTAATCTGAAAGCAACCCGCTGATCCCCCAATTA	480	
DB	421	TTTTTGAAGAAAGCGGACATATTAATCTGAAAGCAACCCGCTGATCCCCCAATTA	480	
QY	481	CAGTATGCTTTCAAGACAAATAATCACTTATCTGTATCATGAAATATCAGCCTGAGGG	540	
DB	481	CAGTATGCTTTCAAGACAAATAATCACTTATCTGTATCATGAAATATCAGCCTGAGGG	540	
QY	541	GACTTGCTGCTACCTTTGAATATGATATGAGACCAAGTTAGTGAATATCAGTTACAGTTT	600	
DB	541	GACTTGCTGCTACCTTTGAATATGATATGAGACCAAGTTAGTGAATATCAGTTACAGTTT	600	
QY	601	TACCTAGCTGAGCTGATTTTGGCTGTTCAAGCCTTCACTGATGGGATACGTCATCGA	660	
DB	601	TACCTAGCTGAGCTGATTTTGGCTGTTCAAGCCTTCACTGATGGGATACGTCATCGA	660	
QY	661	GACATCAAGCCTGAGAACATTTCTGTTGACCGACAGAGACATCAAGCTGTGGATTTT	720	
DB	661	GACATCAAGCCTGAGAACATTTCTGTTGACCGACAGAGACATCAAGCTGTGGATTTT	720	
QY	721	GGAATCTGCGCGGAAATGAAATTCAAACAA--GGTGAATGCAAACTCCGATTTGGGAGC	777	
DB	721	GGAATCTGCGCGGAAATGAAATTCAAACAA--GGTGAATGCAAACTCCGATTTGGGAGC	777	
QY	781	CCAGATTACATGAGCTCTCGAAGTCTGATGTAAGCGGGATGGAAGAAAGCACTTAC	840	
DB	781	CCAGATTACATGAGCTCTCGAAGTCTGATGTAAGCGGGATGGAAGAAAGCACTTAC	840	
QY	841	GAGCTGAGCTGTGACTGTGTGCTGAGTGGGCTGATTTGCTATGAGATATTAATGGGAGA	900	
DB	841	GAGCTGAGCTGTGACTGTGTGCTGAGTGGGCTGATTTGCTATGAGATATTAATGGGAGA	900	
QY	898	TCCCCCTTGGCAGAGGGAACTCTGCGCAAACTTCAATATCAATTAATATTTCCAGCG	960	
DB	898	TCCCCCTTGGCAGAGGGAACTCTGCGCAAACTTCAATATCAATTAATATTTCCAGCG	957	

QY	961	TTTTTGAATTTTCAGATGATCCCAAGTAGACGTACTTTGATCTGATTCAAAAGC	1020
Db	958	TTTTTGAATTTTCAGATGATCCCAAGTAGACGTACTTTGATCTGATTCAAAAGC	1017
QY	1021	TTGTTGTGGGCGACAAAAGAGAGATGAAAGTTTGAAGGTTTGGTGTGCAATCCCTTCTTC	1088
Db	1018	TTGTTGTGGGCGACAAAAGAGAGATGAAAGTTTGAAGGTTTGGTGTGCAATCCCTTCTTC	1077
QY	1081	TCATAAATTGACTGGAACAACATTTCTGTAATCTCTCCGCCCTTGTTCCCAACCTCAAG	1148
Db	1078	TCATAAATTGACTGGAACAACATTTCTGTAAGCTCTCCGCCCTTGTTCCCAACCTCAAG	1137
QY	1141	TCGTACGATGACACTCCAAATTTTGATGAACCAAGAAAGAAATTCGTGGGTTTCACTCTCT	1200
Db	1138	TCGTACGATGACACTCCCAATTTTGATGAACCAAGAAAGAAATTCGTGGGTTTCACTCTCT	1197
QY	1201	CGGTGCGACGTGAGGCCCTCCAGGCTTCCGGGTGAAGAACTGCCGTTTGGGGGTTTTTCG	1266
Db	1198	CGGTGCGACGTGAGGCCCTCCAGGCTTCCGGGTGAAGAACTGCCGTTTGGGGGTTTTTCG	1257
QY	1261	TACAGCAAGGCACTGGGGATTCTTGTTGATCTGTGATCTGTGTGTCGGGCTGGAATCC	1328
Db	1258	TACAGCAAGGCACTGGGGATTCTTGTTGATCTGTGATCTGTGTGTCGGGCTGGAATCC	1317
QY	1321	CCGTGCAAGACTAGCTCCATGTAAGAACTTCTCATCAAAAGCAAAAGCTTCAAGAC	1380
Db	1318	CCGTGCAAGACTAGCTCCATGTAAGAAACTTCTCATCAAAAGCAAAAGCTTCAAGAC	1377
QY	1381	TCTCAGGACAACTGTCAACAATGAGACAGAAATGACCCGGTTACATCCGAGAGTGTCA	1440
Db	1378	TCTCAGGACAAAGTGTCAACAATGAGAGAGAAATGACCCGGTTACATCCGAGAGTGTCA	1437
QY	1441	GAGGTGAGGCTGTGCTTAATGTCAGAAAGAGGTGAGGCTGAAAGGCTCTGAGACTCAGAGA	1500
Db	1438	GAGGTGAGGCTGTGCTTAATGTCAGAAAGAGGTGAGGCTGAAAGGCTCTGAGACTCAGAGA	1497
QY	1501	TCCCTCTCGAGACGAACCTTGCTACTCAATCAACAAGATGCAAGTACTTAAACGAAGT	1560
Db	1498	TCCCTCTCGAGACGAACCTTGCTACTCAATCAACAAGATGCAAGTACTTAAACGAAGT	1557
QY	1551	TTTGAAGCAGACCGGATGAGAGTGTCCCAAGAGATGACAAAGCATCTGCAGCTTCTCAT	1620
Db	1558	TTTGAAGCAGACCGGATGAGAGTGTCCCAAGAGATGACAAAGCATCTGCAGCTTCTCAT	1617
QY	1621	GATATCAAGAGCAGAGCCGGAAGCTTCCAAAGAAATCAAAAGAGAGATGACCGGCTCAAC	1680
Db	1618	GATATCAAGAGCAGAGCCGGAAGCTTCCAAAGAAATCAAAAGAGAGATGACCGGCTCAAC	1677
QY	1681	GTGGAAGAAATGAGGTTGATGATGAATCAGTTTGGAAGAGATCTTGTCTCAGCAAGAGA	1740
Db	1678	GTGGAAGAAATGAGGTTGATGATGAATCAGTTTGGAAGAGATCTTGTCTCAGCAAGAGA	1737
QY	1741	CGAGTGATCTCTACGAATCTGAGCTGAGAGAGTCTTCGCTTGCTGTAAGATTTCAAG	1800
Db	1738	CGAGTGATCTCTACGAATCTGAGCTGAGAGAGTCTTCGCTTGCTGTAAGATTTCAAG	1797
QY	1801	CGGAAACCGACAGAAATTCAGCATTAACCTTTGAAGGCTTAAGGATCAAGGGGAAGCCTGTAA	1860
Db	1798	CGGAAACCGACAGAAATTCAGCATTAACCTTTGAAGGCTTAAGGATCAAGGGGAAGCCTGTAA	1857
QY	1861	GTGGAGAAATATGCGGAATCTGGAGAAAGATCAATAGCTGAGCAGAGCTCAAAATTCAGGAG	1920
Db	1858	GTGGAGAAATATGCGGAATCTGGAGAAAGATCAATAGCTGAGCAGAGCTCAAAATTCAGGAG	1917
QY	1921	CTCCAAAGAAACTGAGAAAGGCTGTAAAGCCAGACACGAGGCGCAACGAGCTGCTGCAAG	1980
Db	1918	CTCCAAAGAAACTGAGAAAGGCTGTAAAGCCAGACACGAGGCGCAACGAGCTGCTGCAAG	1977
QY	1981	AATATCCGCAAGGCAAGGACCGACCGAGAGGAGAGCTGGAAGGCTCGAAGACCGAGAG	2040
Db	1978	AATATCCGCAAGGCAAGGACCGACCGAGAGGAGAGCTGGAAGGCTCGAAGACCGAGAG	2037

QY 2041 GATTCCTCTGAAGCATCAGAAAAGCTGTGTGAAGCTGAGAGACCCGCTATTCTCTG 2100
|
|
|
Db 2038 GATTCCTCTGAAGCATCAGAAAAGCTGTGTGAAGCTGAGAGACCCGCTATTCTCTG 2097
|
|
|
QY 2101 GAGAACAAAGTTAAAGACTAGACCATGAGGCTGAGAAAACAGACTGAAGATGAC 2160
|
|
|
|
Db 2098 GAGAACAAAGTTAAAGACTAGACCATGAGGCTGAGAAAACAGACTGAAGATGAC 2157
|
|
|
|
QY 2161 ATCCAGACAAATATCCCAACGATCCAGACAGTGGCTGATTAATTTCTGAGCTCGAAGAG 2220
|
|
|
|
Db 2158 ATCCAGACAAATATCCCAACGATCCAGACAGTGGCTGATTAATTTCTGAGCTCGAAGAG 2217
|
|
|
|
QY 2221 AAACATGTGGAGGCTCAAGTCTCAGCCCAAGCACTTAAGTGCACCTGAAAACAGAAAGAG 2280
|
|
|
|
Db 2218 AAACATGTGGAGGCTCAAGTCTCAGCCCAAGCACTTAAGTGCACCTGAAAACAGAAAGAG 2277
|
|
|
|
QY 2281 CAGACATAGAGAAAAGATTAAAGTGTGAGCAATGAGATTAAGAAAAGCCTGGCTGAC 2340
|
|
|
|
Db 2278 CAGACATAGAGAAAAGATTAAAGTGTGAGCAATGAGATTAAGAAAAGCCTGGCTGAC 2337
|
|
|
|
QY 2341 AAGGAGACACTGAGAAACATGATGCAGAGACACGAGAGAGAGGCCCATGAGAGAGGCAAA 2400
|
|
|
|
Db 2338 AAGGAGACACTGAGAAACATGATGCAGAGACACGAGAGAGAGGCCCATGAGAGAGGCAAA 2397
|
|
|
|
QY 2401 ATTCTCAGCAACAGAAAGCGATGATCAATGCTATGGAATTCGAAGATCAAGTCCCTGAAA 2460
|
|
|
|
Db 2398 ATTCTCAGCAACAGAAAGCGATGATCAATGCTATGGAATTCGAAGATCAAGTCCCTGAAA 2457
|
|
|
|
QY 2461 CAGAGATTGTGGAATCTGTCTGAAGCCCAATTAATTTGACAGAAATAGCAGCTTTTAAAC 2520
|
|
|
|
Db 2458 CAGAGATTGTGGAATCTGTCTGAAGCCCAATTAATTTGACAGAAATAGCAGCTTTTAAAC 2517
|
|
|
|
QY 2521 CAAAGGAACATGAAAGGCCCAAGAAAGATGATTTCTGAACCTCAGGCAACAGAAATTTTAC 2580
|
|
|
|
Db 2518 CAAAGGAACATGAAAGGCCCAAGAAAGATGATTTCTGAACCTCAGGCAACAGAAATTTTAC 2577
|
|
|
|
QY 2581 CTGAGACACAGGCTGGGAAAGTTGAGAGCCCAAGAACCTGAGAGACAGCTGAG 2640
|
|
|
|
Db 2578 CTGAGACACAGGCTGGGAAAGTTGAGAGCCCAAGAACCTGAGAGACAGCTGAG 2637
|
|
|
|
QY 2641 AAGATCGCCACCAAGACCAAGTGAACAAGATGGCTGCTGGAACCTGAGACAAAGTTG 2700
|
|
|
|
Db 2638 AAGATCGCCACCAAGACCAAGTGAACAAGATGGCTGCTGGAACCTGAGACAAAGTTG 2697
|
|
|
|
QY 2701 CGGAGGCTCAGTCTAGAGACAGAGACAGAACTGAGCTCAAGCCCAAGCTCAGAG 2760
|
|
|
|
Db 2698 CGGAGGCTCAGTCTAGAGACAGAGACAGAACTGAGCTCAAGCCCAAGCTCAGAG 2757
|
|
|
|
QY 2761 CTACAGCTCTCCCTGAGAGAGCGGAGTCAAGTTGACAGCCCTGCAAGCTGCA 2820
|
|
|
|
Db 2758 CTACAGCTCTCCCTGAGAGAGCGGAGTCAAGTTGACAGCCCTGCAAGCTGCA 2817
|
|
|
|
QY 2821 GCCCTGAGAGCCAGCTTCCGCAAGGCAAGACAGAGCTGGAAGACACACAGAGAGCT 2880
|
|
|
|
Db 2818 GCCCTGAGAGCCAGCTTCCGCAAGGCAAGAGCTGGAAGACACACAGAGAGCT 2877
|
|
|
|
QY 2881 GAGAGGAGATCCAGGCACTCAGGGCAATAGAGTGAATCCAGCCCAATTTGATGCT 2940
|
|
|
|
Db 2878 GAGAGGAGATCCAGGCACTCAGGGCAATAGAGTGAATCCAGCCCAATTTGATGCT 2937
|
|
|
|
QY 2941 CTTCTGAACAGCTGTACTGTATTCACAGACTGTGAGAGAGCACTTAACCAAGTGAAGCCGAG 3000
|
|
|
|
Db 2938 CTTCTGAACAGCTGTACTGTATTCACAGACTGTGAGAGAGCACTTAACCAAGTGAAGCCGAG 2997
|
|
|
|
QY 3001 GACAAAGCTGAACTCAACCAAACTTTCTATTGTTCCAACCAACTGCATGAGGCTTCT 3060
|
|
|
|
Db 2998 GACAAAGCTGAACTCAACCAAACTTTCTATTGTTCCAACCAACTGCATGAGGCTTCT 3057
|
|
|
|
QY 3061 GGGCCCAACGACGAGATTGTACAACTGGAAGTGAAGCACTCCGCGGGGAGATC 3120
|
|
|
|
Db 3058 GGGCCCAACGACGAGATTGTACAACTGGAAGTGAAGCACTCCGCGGGGAGATC 3117
|
|
|
|
QY 3121 ACGGAAAGAGATGACGCTTACCAAGCAGAAAGCAATGAGAGCTCTGAAAGACACG 3180
|
|
|
|

Db 3118 ACGGAAAGAGATGACGCTTACCAAGCAGAAAGCAATGAGAGCTCTGAAAGACACG 3177
|
|
|
|
QY 3181 TGCAACATGCTGAGAGAAACAGGTCATGATTTGGAGGCCCTTAAACGATGAGCTGTAGAA 3240
|
|
|
|
Db 3178 TGCAACATGCTGAGAGAAACAGGTCATGATTTGGAGGCCCTTAAACGATGAGCTGTAGAA 3237
|
|
|
|
QY 3241 AAAGAGCGAGATGAGAGGCCCTGAGAGAGCGTCTGGGTGATGAGAAATCCCAAGTTGAG 3300
|
|
|
|
Db 3238 AAAGAGCGAGATGAGAGGCCCTGAGAGAGCGTCTGGGTGATGAGAAATCCCAAGTTGAG 3297
|
|
|
|
QY 3301 TGTGGGTTGAGAGCTGCAAGATGCTGGAACCCAGAAACAGACAGGCGAGAGGCC 3360
|
|
|
|
Db 3298 TGTGGGTTGAGAGCTGCAAGATGCTGGAACCCAGAAACAGACAGGCGAGAGGCC 3357
|
|
|
|
QY 3361 GATCAGCGGATCAACCGAGTCTCGCCAGAGTGTGAGAGCTGCGAGTGAAGAGACCAAGGCT 3420
|
|
|
|
Db 3358 GATCAGCGGATCAACCGAGTCTCGCCAGAGTGTGAGAGCTGCGAGTGAAGAGACCAAGGCT 3417
|
|
|
|
QY 3421 GAGATTCTGCTCTGACAGAGGCTCTCAAAAGACGAAAGCTGAAGGCCGAGAGGCTTCT 3480
|
|
|
|
Db 3418 GAGATTCTGCTCTGACAGAGGCTCTCAAAAGACGAAAGCTGAAGGCCGAGAGGCTTCT 3477
|
|
|
|
QY 3481 GACAAAGTCAATGACCTGAGAGAGAAACATGCTATGCTTGAATGAATGCCGAAAGCTTA 3540
|
|
|
|
Db 3478 GACAAAGTCAATGACCTGAGAGAGAAACATGCTATGCTTGAATGAATGCCGAAAGCTTA 3537
|
|
|
|
QY 3541 CAGCAGAGCTGGAGACTGAACGAGAGCTCAACAGAGGCTCTGGAAGAGCAAGCCAA 3600
|
|
|
|
Db 3538 CAGCAGAGCTGGAGACTGAACGAGAGCTCAACAGAGGCTCTGGAAGAGCAAGCCAA 3597
|
|
|
|
QY 3601 TTACAGCAGAGATGAGACTGTCAGAAAAATCACATTTTCGTGACTGACTAAGGACTGCA 3660
|
|
|
|
Db 3598 TTACAGCAGAGATGAGACTGTCAGAAAAATCACATTTTCGTGACTGACTAAGGACTGCA 3657
|
|
|
|
QY 3661 GAAGCTTAAGTGGGCTGATCTATGTAAGACAGAAAGAGTGAAGTCTGAGATCAGCTG 3720
|
|
|
|
Db 3658 GAAGCTTAAGTGGGCTGATCTATGTAAGACAGAAAGAGTGAAGTCTGAGATCAGCTG 3717
|
|
|
|
QY 3721 GAAACATTAAGGTTCTTAATTCATGAAGAGTGAAGTGAAGGCACTAATTTCTCA 3780
|
|
|
|
Db 3718 GAAACATTAAGGTTCTTAATTCATGAAGAGTGAAGTGAAGGCACTAATTTCTCA 3777
|
|
|
|
QY 3781 CAAACCAACTCATTTGATTTTCTGCAAGCCCAATGAGCAACCTGTAAAGAAAAAG 3840
|
|
|
|
Db 3778 CAAACCAACTCATTTGATTTTCTGCAAGCCCAATGAGCAACCTGTAAAGAAAAAG 3837
|
|
|
|
QY 3841 GTTCTCTGCAATCAATGAGCTGAGAGCTGAGAGAGAGAAAGCTCGCTGTGCA 3900
|
|
|
|
Db 3838 GTTCTCTGCAATCAATGAGCTGAGAGCTGAGAGAGAGAAAGCTCGCTGTGCA 3897
|
|
|
|
QY 3901 GAGCTAAGAGAAAGCCCTTCAAGAAAGCCGCAATGAGTCCGCTCCGCGGAGAGAGCT 3960
|
|
|
|
Db 3898 GAGCTAAGAGAAAGCCCTTCAAGAAAGCCGCAATGAGTCCGCTCCGCGGAGAGAGCT 3957
|
|
|
|
QY 3961 GCCCACCAGCAAGAGAGAGCAACCCACACCACTCAAGCCAGAGCCAGAGGCAAGAG 4020
|
|
|
|
Db 3958 GCCCACCAGCAAGAGAGAGAGCAACCCACACCACTCAAGCCAGAGCCAGAGGCAAGAG 4017
|
|
|
|
QY 4021 ATGCGCATGTCGCGCATGTCGCGGTGCGCAGAGACCAAGCCCAAGTGCATGAGCTGCTG 4080
|
|
|
|
Db 4018 ATGCGCATGTCGCGCATGTCGCGGTGCGCAGAGACCAAGCCCAAGTGCATGAGCTGCTG 4077
|
|
|
|
QY 4081 GCCCGCGCATCCAGCGCGAGAAAGAGATCTTCAACTCAGAGAAATTAAGTGGCGCTT 4140
|
|
|
|
Db 4078 GCCCGCGCATCCAGCGCGAGAAAGAGATCTTCAACTCAGAGAAATTAAGTGGCGCTT 4137
|
|
|
|
QY 4141 AAGGAAGCATGACCAACAATATTTCTCAAGGATTCACAGTGAAGCTGAACATGAGAGCC 4200
|
|
|
|
Db 4138 AAGGAAGCATGACCAACAATATTTCTCAAGGATTCACAGTGAAGCTGAACATGAGAGCC 4197
|
|
|
|
QY 4201 ACAAGTGTCTGTGTCTGTGATACCGTGCACTTTGAGAGCCAGGCAATCCAAATGTCTC 4260
|
|
|
|

Db	4198	ACAAAGTGTGCTGTGTGTGTGGAATACCGTGACCTTGGACCGCAGGATCCAAATGTCTA	4257
QY	4261	GAATGTCAAGTGAATGTGTACCCCAAGTGCTCAAGTCTGTGGCCAGCCACCTGGCGCTTG	4320
Db	4258	GAATGTCAAGTGTATGTGTACCCCAAGTGCTCAAGTGTGGCCAGCCACCTGGCGCTTG	4317
QY	4321	CCGTGCTGAATATATGGCAACACTTTCACCGAAGGCTTCTGCGGTGACAAATGAATCTCCCA	4380
Db	4318	CCGTGCTGAATATATGGCAACACTTTCACCGAAGGCTTCTGCGGTGACAAATGAATCTCCCA	4377
QY	4381	GGTCTTCAGACCAAGAGCCCGACAGACGCTTGACCTGGAAAGGTTGATGAAGTGGCC	4440
Db	4378	GGTCTTCAGACCAAGAGCCCGACAGACGCTTGACCTGGAAAGGTTGATGAAGTGGCC	4437
QY	4441	AGGAATTAACAACGAGGACGAGAAGCTGGGACAGAAAGTCAATTTGCTTGAAGAGATCA	4500
Db	4438	AGGAATTAACAACGAGGACGAGAAGCTGGGACAGAAAGTCAATTTGCTTGAAGAGATCA	4497
QY	4501	AAAGTCTTCATTTATATGACATATGAACCGAGAAAGCTGGAACAGAGGCGGTGAAAGATT	4560
Db	4498	AAAGTCTTCATTTATATGACATATGAACCGAGAAAGCTGGAACAGAGGCGGTGAAAGATT	4557
QY	4561	GAGCTGTGCTTCCCGGACGGGGATGTATCTATTCATGGTGGCGTTGGTGGCTTCCGAATC	4620
Db	4558	GAGCTGTGCTTCCCGGACGGGGATGTATCTATTCATGGTGGCGTTGGTGGCTTCCGAATC	4617
QY	4621	GCAAATACAGCCAAAGACGATGTGCCATACATCTGAAGATGATGAATCTCACCCGACACC	4680
Db	4618	GCAAATACAGCCAAAGACGATGTGCCATACATCTGAAGATGATGAATCTCACCCGACACC	4677
QY	4681	ACCTGTCTGGCCCGGGAGAAACCTCTATCTTGTGTAAGTCTCCAGCTTCCGTGACAAACAGCC	4740
Db	4678	ACCTGTCTGGCCCGGGAGAAACCTCTATCTTGTGTAAGTCTCCAGCTTCCGTGACAAACAGCC	4737
QY	4741	TGGGTCAACCGCTTGAATCAAGTTGTGACAGTGGGAGAGTTTCTAGGAAAGAAAGCAGAA	4800
Db	4738	TGGGTCAACCGCTTGAATCAAGTTGTGACAGTGGGAGAGTTTCTAGGAAAGAAAGCAGAA	4797
QY	4801	GCTGATGCTAAATCTGCTTGAAGACTCCCTGTCTGAACCTGAGAGGTGATGACCCGTATGAC	4860
Db	4798	GCTGATGCTAAATCTGCTTGAAGACTCCCTGTCTGAACCTGAGAGGTGATGACCCGTATGAC	4857
QY	4861	ATGAATCTGACAGCTGCCCTTCAATGACCAAGTGTGTGGGACCTGAGAAAGGCGCTC	4920
Db	4858	ATGAATCTGACAGCTGCCCTTCAATGACCAAGTGTGTGGGACCTGAGAAAGGCGCTC	4917
QY	4921	TACGCCCTGAAATGTCTGAAAGAACTCCCTAACCCATGTCCAGGAATTTGAGACAGCTTTC	4980
Db	4918	TACGCCCTGAAATGTCTGAAAGAACTCCCTAACCCATGTCCAGGAATTTGAGACAGCTTTC	4977
QY	4981	CAAAATTAATATATCAAGACCTGAGAGAGCTACTCATATGATGACGAGAGAGAGCGGCA	5040
Db	4978	CAAAATTAATATATCAAGACCTGAGAGAGCTACTCATATGATGACGAGAGAGAGCGGCA	5037
QY	5041	CTGTGTCTTGTGGACGTGSAAGAAAGTGAAACAGTCCCTGGGCCAGTCCACCTGCTGCC	5100
Db	5038	CTGTGTCTTGTGGACGTGSAAGAAAGTGAAACAGTCCCTGGGCCAGTCCACCTGCTGCC	5097
QY	5101	CAGCCCGGACATCTACCCCAATTTTGAAGCTGTCAAGGGCTGGCCACTTGTGGGGGCA	5160
Db	5098	CAGCCCGGACATCTACCCCAATTTTGAAGCTGTCAAGGGCTGGCCACTTGTGGGGGCA	5157
QY	5161	GGCAAGATTGAAACGGGCTCTGCATCTGTGACAGCCATGGCCAGCAAAAGTGTCAATCTTC	5220
Db	5158	GGCAAGATTGAAACGGGCTCTGCATCTGTGACAGCCATGGCCAGCAAAAGTGTCAATCTTC	5217
QY	5221	CGCTACAAACGAAACCTCAGCAATCTGCATCCGGAAGAGATGAGACCTCAGAGCCC	5280
Db	5218	CGCTACAAACGAAACCTCAGCAATCTGCATCCGGAAGAGATGAGACCTCAGAGCCC	5277
QY	5281	TGCAAGCTGTATCCCACTTCAACCAATTAACAGTATCTCTATTGAAACCAATTAATTTCAAGAA	5340
Db	5278	TGCAAGCTGTATCCCACTTCAACCAATTAACAGTATCTCTATTGAAACCAATTAATTTCAAGAA	5337

QY	5341	ATCGACATGAAGACAGTAAACAAGCTCGAGGAATTTCTCGAGTAAGAAATGACCAATTCCTTGGCA	5400
Db	5338	ATCGACATGAAGACAGTAAACAAGCTCGAGGAATTTCTCGAGTAAGAAATGACCAATTCCTTGGCA	5397
QY	5401	CTGTGCTGTGTTGGCGCGCTCTTCCCAAAGCTTCCCTGTCTCAATCGAGAGTGAACAGC	5460
Db	5398	CTGTGCTGTGTTGGCGCGCTCTTCCCAAAGCTTCCCTGTCTCAATCGAGAGTGAACAGC	5457
QY	5461	GCAGGGCAGCGAGAGAGTACTTGCTGTGTTCCAAGATTGGAGTTCGTGAGATTCT	5520
Db	5458	GCAGGGCAGCGAGAGAGTACTTGCTGTGTTCCAAGATTGGAGTTCGTGAGATTCT	5517
QY	5521	TACGGAAACGTAAGCCGACAGAGATCTCAAGGGAGTGGCTTAACCTTTGGCCTTTGGC	5580
Db	5518	TACGGAAACGTAAGCCGACAGAGATCTCAAGGGAGTGGCTTAACCTTTGGCCTTTGGC	5577
QY	5581	TACAGAGAACCCTATCTGTGTTGTGACCCACTTCAACTCACTCGAAGTAATTGAGATCCAG	5640
Db	5578	TACAGAGAACCCTATCTGTGTTGTGACCCACTTCAACTCACTCGAAGTAATTGAGATCCAG	5637
QY	5661	GCAAGCTTCTCAAGAGGGAATCCCTGGCCGAGCGTAACCTGGAATATCCGAATCCGCGCTAC	5700
Db	5638	GCACAGCTTCTCAAGAGGGAATCCCTGGCCGAGCGTAACCTGGAATATCCGAATCCGCGCTAC	5697
QY	5701	CTGGGCCCCGTGCATTTCTCTGAGAGCGATTACTTGGCGTCTCATATCAAGGATTAATTA	5760
Db	5688	CTGGGCCCCGTGCATTTCTCTGAGAGCGATTACTTGGCGTCTCATATCAAGGATTAATTA	5757
QY	5761	AGGTCATTTTCTGCAAGGGAACCTCGTGAAGGAGTCCGGCACTGAACAACAACCGGGGC	5820
Db	5758	AGGTCATTTTCTGCAAGGGAACCTCGTGAAGGAGTCCGGCACTGAACAACAACCGGGGC	5817
QY	5821	CCGTCCACTTCCCGCAGAGATCCCAACAAGCAGAGGCCCACTCAAGTACAACAGACATCT	5880
Db	5818	CCGTCCACTTCCCGCAGAGATCCCAACAAGCAGAGGCCCACTCAAGTACAACAGACATCT	5877
QY	5881	ACCAAGGCGTGGGCTCTCAAGCCGACGCGCCGCGGAAAGGCCCAACCGCGACAGAGCCA	5940
Db	5878	ACCAAGGCGTGGGCTCTCAAGCCGACGCGCCGCGGAAAGGCCCAACCGCGACAGAGCCA	5937
QY	5941	AGCACACCCCAACCGCTACCGCGAGGGCGGACCGAGCTGGCGAGGGAACAAGTCTCTGGC	6000
Db	5938	AGCACACCCCAACCGCTACCGCGAGGGCGGAGCTGGCGAGGGAACAAGTCTCTGGC	5997
QY	6001	CGCCCCCTTGGAGCGACAAGATCCCCCTGGCCGGATATCTCAAGCAACCGGAGAGACCGGTCC	6060
Db	5998	CGCCCCCTTGGAGCGACAAGATCCCCCTGGCCGGATATCTCAAGCAACCGGAGAGACCGGTCC	6057
QY	6061	CCCGCAGAGGCTGTTTGAAGAAGCAGCAGAGGAGGCGCGGCTGCTGCGGGAGCCGTGAGAGCC	6120
Db	6058	CCCGCAGAGGCTGTTTGAAGAAGCAGCAGAGGAGGCGCGGCTGCTGCGGGAGCCGTGAGAGCC	6117
QY	6121	CCGCTGTCCACAGGTGAACAAGGTCTGGGACCAAGTCTCTGGC	6155
Db	6118	CCGCTGTCCACAGGTGAACAAGGTGAAGGACCAAGTCTCTGGC	6152
RESULT 10			
LOCUS	AY257469	8576 bp	mRNA linear PRI 23-APR-2003
DEFINITION	Homo sapiens rho/rac-interacting citron kinase (CIT) mRNA, complete cds.		
ACCESSION	AY257469		
VERSION	AY257469.1	GI:30088969	
KEYWORDS			
SOURCE			
ORGANISM	Homo sapiens (human)		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.		
AUTHORS	Huang,C.Q., Wu,S.T., Shan,Y.X., Liu,S. and Xiao,P.J.		
TITLE	Direct Submission		

QY 1261 TACAGCAAGGCACTGGGGATTCTGTAGATCTGAGTCTGTTGTGTGGGCTTGAAGCTCC 1320
DB 1314 TACAGCAAGGCACTGGGGATTCTGTGATCTGATGCTGTGTGTGTGGGCTTGAAGCTCC 1373
QY 1321 CTGCGCAAGCACTAGCTCCATGGAAGAACTTCTCATCAAAAGCAAGAGTCAAGAC 1380
DB 1374 CTGCGCAAGCACTAGCTCCATGGAAGAACTTCTCATCAAAAGCAAGAGTCAAGAC 1433
QY 1381 TCTCAGGACAAAGTGTCACAAGATGGAAGCAAGAAATGACCCGGTTACATCCGAGTGTCA 1440
DB 1434 TCTCAGGACAAAGTGTCACAAGATGGAAGCAAGAAATGACCCGGTTACATCCGAGTGTCA 1493
QY 1441 GAGGTGAGGCTGTGCTTACTGCAAGAGAGGTGAGCTGAAAGGCTTGAAGCTCAGAGA 1500
DB 1494 GAGGTGAGGCTGTGCTTACTGCAAGAGAGGTGAGCTGAAAGGCTTGAAGCTCAGAGA 1553
QY 1501 TCCCTCTGAGAGCAGAGCTTTGCTACTACATCA CAGATGCAAGTACCTTAAAGCGAAGT 1560
DB 1554 TCCCTCTGAGAGCAGAGCTTTGCTACTACATCA CAGATGCAAGTACCTTAAAGCGAAGT 1613
QY 1561 TTGAGAGAACCAAGGATGAGAGTGTCCAGAGAGATGACAAAGCACTGCAAGCTTCTCAT 1620
DB 1614 TTGAGAGAACCAAGGATGAGAGTGTCCAGAGAGATGACAAAGCACTGCAAGCTTCTCAT 1673
QY 1621 GATATCAGAGAGCAGAGCCGGAAGCTCCAAGAAATCAAAAGAGAGAGTACCAAGGCTCAA 1680
DB 1674 GATATCAGAGAGCAGAGCCGGAAGCTCCAAGAAATCAAAAGAGAGAGTACCAAGGCTCAA 1733
QY 1681 GTGGAAGAAATGAGGTGATGATGAATCAGTTGGAAGAGAGTCTGTCTCAGCAAGAGA 1740
DB 1734 GTGGAAGAAATGAGGTGATGATGAATCAGTTGGAAGAGAGTCTGTCTCAGCAAGAGA 1793
QY 1741 CGAGAGTACTCTTACCAATCTGAGCTGAGAGAGTCTTGCTGTGGAAGATTCAG 1800
DB 1794 CGAGAGTACTCTTACCAATCTGAGAGAGTCTTGCTGTGGAAGATTCAG 1853
QY 1801 CGGAAGAGCAGAGATGTGAGCATTAATCTTTGAAGCTTAAGATCAAGGAGAGCTGAA 1860
DB 1854 CGGAAGAGCAGAGATGTGAGCATTAATCTTTGAAGCTTAAGATCAAGGAGAGCTGAA 1913
QY 1861 GTGGAGAAATATGCGAACTGAGAGAGATCAATGCTGAGCAGAGCTCAAAATTCAGAG 1920
DB 1914 GTGGAGAAATATGCGAACTGAGAGAGATCAATGCTGAGCAGAGCTCAAAATTCAGAG 1973
QY 1921 CTCGAAGAAATCTGAGAGAGGCTGTTAAAGCCAGACCGAGCTGCTGCGAG 1980
DB 1974 CTCGAAGAAATCTGAGAGAGGCTGTTAAAGCCAGACCGAGCTGCTGCGAG 2033
QY 1981 AATATCCGCGAGGCAAGAGCGAGCGAGAGGAGCTGGAAGAGCTGCAGAACCGAGAG 2040
DB 2034 AATATCCGCGAGGCAAGAGCGAGCGAGAGGAGCTGGAAGAGCTGCAGAACCGAGAG 2093
QY 2041 GATTCTTCTGAAGGATCAGAAAGAGCTGTGGAAGCTGAGAGAGCGCCGCAATTCCTG 2100
DB 2094 GATTCTTCTGAAGGATCAGAAAGAGCTGTGGAAGCTGAGAGAGCGCCGCAATTCCTG 2134
QY 2101 GAGAACAGGTTAAGAGCTAGAGCCATGAGAGCTAGAGAAAAAGACTGAAGATGAC 2160
DB 2135 ----- 2134
QY 2161 ATCCAGACAAATATCCCAACAGATCCAGATGAGCTGATTAATTTCTGAGAGCTCGAGAG 2220
DB 2135 -----GAGAGCTCGAGAGAG 2147
QY 2221 AAACATCGGAGGCGCCAAATCTCAGCCAGACACTTGAAGTCACTGAAAGAGAGAG 2280
DB 2148 AAACATCGGAGGCGCCAAATCTCAGCCAGACACTTGAAGTCACTGAAAGAGAGAG 2207
QY 2281 CAGCACTATGAGAGAAAGATTAAAGTGTGGAACAATCAGATTAAGAGAGAGAGCTGCTGAC 2340
DB 2208 CAGCACTATGAGAGAAAGATTAAAGTGTGGAACAATCAGATTAAGAGAGAGAGCTGCTGAC 2267
QY 2341 AAGAGACACTGAGAGAACATGATGACAGAGACAGAGAGAGAGGCCCAATGAGAGAGGCAAA 2400

DB 2268 AAGGAGCACTGGAGAACATGATGACAGACAGAGAGAGAGGCCCATGAGAGAGGCAAA 2327
QY 2401 ATTCTCAGCAACAGAAAGGAGATGATCAATGCTATGGAATTCAGAGATCAGATCCCTGAAA 2460
DB 2328 ATTCTCAGCAACAGAAAGGAGATGATCAATGCTATGGAATTCAGAGATCAGATCCCTGAAA 2387
QY 2461 CAGAGATTTGGAACTGTCTGAAAGCCATTAACCTTGACAGAAATGACAGTCTTTTAAAC 2520
DB 2388 CAGAGATTTGGAACTGTCTGAAAGCCATTAACCTTGACAGAAATGACAGTCTTTTAAAC 2447
QY 2521 CAAGAGAACATGAAAGGCCCAAGAGAGATGATTTCTGAATCTCAGAGGCAAGAAATTTTAC 2580
DB 2448 CAAGAGAACATGAAAGGCCCAAGAGAGATGATTTCTGAATCTCAGAGGCAAGAAATTTTAC 2507
QY 2581 CTGGAACACAGGCTGGGAAGTTGAGAGGCCAGAACCGAATACTGAGAGAGAGCTGAGAG 2640
DB 2508 CTGGAACACAGGCTGGGAAGTTGAGAGGCCAGAACCGAATACTGAGAGAGAGCTGAGAG 2567
QY 2641 AAGATCAGCCACCAAGACCA CAGTGA CAAAGATCGGCTGCTGGAATCTGAGACAGAAATTG 2700
DB 2568 AAGATCAGCCACCAAGACCA CAGTGA CAAAGATCGGCTGCTGGAATCTGAGACAGAAATTG 2627
QY 2701 CGGAGGTGAGTGTAGGACAGAGAGCAAGAACTGAGAGCTCAAGGCGCAGCTCAGAGAG 2760
DB 2628 CGGAGGTGAGTGTAGGACAGAGAGCAAGAACTGAGAGCTCAAGGCGCAGCTCAGAGAG 2687
QY 2761 CTACAGCTCTCCCTGAGAGAGCGGAGTCAAGTTGACAGCCCTGAGAGCTGACCGGCG 2820
DB 2688 CTACAGCTCTCCCTGAGAGAGCGGAGTCAAGTTGACAGCCCTGAGAGCTGACCGGCG 2747
QY 2821 GCCCTGGAAGGCCAGCTTTGCCAGGCGGAAGACAGAGCTGGAAGAGACCA CAGCAAGAGCT 2880
DB 2748 GCCCTGGAAGGCCAGCTTTGCCAGGCGGAAGACAGAGCTGGAAGAGACCA CAGCAAGAGCT 2807
QY 2881 GAAGAGAGATCCAGGCACTCACGCGCACATAGAGATGAATCCAGGCGCAATTTGATGCT 2940
DB 2808 GAAGAGAGATCCAGGCACTCACGCGCACATAGAGATGAATCCAGGCGCAATTTGATGCT 2867
QY 2941 CTTCTGTAAGCTGTATCTGTAAATCA CAGACTGGAAGAGCAGCTTAACCGAGCTGACCGAG 3000
DB 2868 CTTCTGTAAGCTGTATCTGTAAATCA CAGACTGGAAGAGCAGCTTAACCGAGCTGACCGAG 2927
QY 3001 GACAGAGCTGAATCTCAACCAAGAACTTTACTTGTCCAAACAACTGATGAGCTTCT 3060
DB 2928 GACAGAGCTGAATCTCAACCAAGAACTTTACTTGTCCAAACAACTGATGAGCTTCT 2987
QY 3061 GGCAGCAACGAGAGAGTTGTACAACCTGCGAAGTGAAGTGAACATCTCCGCGGGAGATC 3120
DB 2988 GGCAGCAACGAGAGAGTTGTACAACCTGCGAAGTGAAGTGAACATCTCCGCGGGAGATC 3047
QY 3121 ACGGAACGAGAGATCAGCTTACAGCCAGAGCAAAAGCATGAGAGGCTTGAAGACCAAG 3180
DB 3048 ACGGAACGAGAGATCAGCTTACAGCCAGAGCAAAAGCATGAGAGGCTTGAAGACCAAG 3107
QY 3181 TGCACCATGCTGAGAGAACAGGCTCATGGAATTTGAGAGCCCTTAACGATGAGCTGTAGAA 3240
DB 3108 TGCACCATGCTGAGAGAACAGGCTCATGGAATTTGAGAGCCCTTAACGATGAGCTGTAGAA 3167
QY 3241 AAAGAGCGGAGTGGAGAGGCTGAGAGAGCGCTCGGGGTGATGAGAAATCCAGATTGAG 3300
DB 3168 AAAGAGCGGAGTGGAGAGGCTGAGAGAGCGCTCGGGGTGATGAGAAATCCAGATTGAG 3227
QY 3301 TGTCCGGTTTCAGAGACTCAGAGAGATGCTGACACCGAGAAACAGAGAGGCGGAGAGCC 3360
DB 3228 TGTCCGGTTTCAGAGACTCAGAGAGATGCTGACACCGAGAAACAGAGAGGCGGAGAGCC 3287
QY 3361 GATCAGCGGATCACAGAGTCTTCCAGAGTGTGAGAGCTTGAGAGTGAAGAGACCAAGGCT 3420
DB 3288 GATCAGCGGATCACAGAGTCTTCCAGAGTGTGAGAGCTTGAGAGTGAAGAGACCAAGGCT 3347
QY 3421 GAGATTTCTGCTCTGACAGAGCTTCTCAAGAGAGAGCTGAAGCCGAGAGCTTCTCT 3480

Db 3348 GAGATTCTCGCTCTGCAAGGCTCTCAAAAGCAGAAAGCTGAAGGCCGAGAGCCTTCT 3407
Qy 3481 GACAGCTCAATGACTCTGAGAAAGACATGCTATGCTTGAAATGAATGCCGAGAGCTTA 3540
Db 3408 GACAGCTCAATGACTCTGAGAAAGACATGCTATGCTTGAAATGAATGCCGAGAGCTTA 3467
Qy 3541 CAGCAGAAAGCTGAGACTGAAACGAGAGCTCAAAAGAGGCTTTGGAAGACAGCCAA 3600
Db 3468 CAGCAGAAAGCTGAGACTGAAACGAGAGCTCAAAAGAGGCTTTGGAAGAGCCAA 3527
Qy 3601 TTACAGAGAGAGAGAGCTGAGAGAAATCAGATTTTCGCTGACTCAAGAGCTGCA 3660
Db 3528 TTACAGAGAGAGAGAGCTGAGAGAAATCAGATTTTCGCTGACTCAAGAGCTGCA 3587
Qy 3661 GAAGCTTAGATCGGGCTGATCTTACTGAGACAGAAAGAGTGACTTGAAGTATCACTG 3720
Db 3588 GAAGCTTAGATCGGGCTGATCTTACTGAGACAGAAAGAGTGACTTGAAGTATCACTG 3647
Qy 3721 GAAAACTTCAAGGTTCTCTATTTCTCATGAAAAGGTGAAAATGGAAGCATATTTCTCA 3780
Db 3648 GAAAACTTCAAGGTTCTCTATTTCTCATGAAAAGGTGAAAATGGAAGCATATTTCTCA 3707
Qy 3781 CAAACCAACTCATGATGTTTTCTGCAAGCCAAATGGAACCACTGCTAAAGAAAAA 3838
Db 3708 CAAACCAACTCATGATGTTTTCTGCAAGCCAAATGGAACCACTGCTAAAGAAAAA 3767
Qy 3839 -----AGTTCTCTGAGTAC 3855
Db 3768 GGTTTATTTAGTGAACGAGAAAGAGACCTGCTTTACCAACAGGTTCTCTGCACTAC 3827
Qy 3856 AATGAGCTGAAGCTGGCTCTGAGAGAGAGAAAGCTGCTGTGTGCAAGCTAGAGAAAGCC 3915
Db 3828 AATGAGCTGAAGCTGGCTCTGAGAGAGAGAAAGCTGCTGTGTGCAAGCTAGAGAAAGCC 3887
Qy 3916 CTTTCAGAAAGCCCGCATGAGCTCCGCTCCGCCCGGAGAGAGAGTGGCCCAACCCGAAAGCA 3975
Db 3888 CTTTCAGAAAGCCCGCATGAGCTCCGCTCCGCCCGGAGAGAGAGTGGCCCAACCCGAAAGCA 3947
Qy 3976 ACCGAGCAACCAACCCATCCACGCGACCCACGCGAGAGAGAGTGCATGCGATGCGCC 4035
Db 3948 ACCGAGCAACCAACCCATCCACGCGACCCACGCGAGAGAGAGTGCATGCGATGCGCC 4007
Qy 4036 ATGTGTGGGTGCGCCAGAGACCAAGCCCATGAGCTGTGCGCCCGCCATCCAGC 4095
Db 4008 ATGTGTGGGTGCGCCAGAGACCAAGCCCATGAGCTGTGCGCCCGCCATCCAGC 4067
Qy 4096 CGCAGAAAGAGTTCCTCACTCCAGAGAAATTTAGTGGGCTTTAAGAAACGATGCAC 4155
Db 4068 CGCAGAAAGAGTTCCTCACTCCAGAGAAATTTAGTGGGCTTTAAGAAACGATGCAC 4127
Qy 4156 CACAAATTTCTCAACGATTCGAAGTGAACAATGAGAGCCAAAGTGTGCTGTG 4215
Db 4128 CACAAATTTCTCAACGATTCGAAGTGAACAATGAGAGCCAAAGTGTGCTGTG 4187
Qy 4216 TGTCTGATACCGTGCATTTGAGAGCCAGGATCCAAATGTCATGATGCAAGTATG 4275
Db 4188 TGTCTGATACCGTGCATTTGAGAGCCAGGATCCAAATGTCATGATGCAAGTATG 4247
Qy 4276 TGTCAACCCCAAGTGTCCAGGTGCTTGCACGCCACTTGGCGCTTGCCTGTAATATGCC 4335
Db 4248 TGTCAACCCCAAGTGTCCAGGTGCTTGCACGCCACTTGGCGCTTGCCTGTAATATGCC 4307
Qy 4336 ACACACTTCAACCGAGGCTTCTGCGGTGACAAATGAACCTCCAGAGTTCACAGCAAG 4395
Db 4308 ACACACTTCAACCGAGGCTTCTGCGGTGACAAATGAACCTCCAGAGTTCACAGCAAG 4367
Qy 4396 GAGCCAGCAGCAGACTTGCACCTTGAAGGAGTGAAGGTGCCAGAAATTAACAAACGA 4455
Db 4368 GAGCCAGCAGCAGACTTGCACCTTGAAGGAGTGAAGGTGCCAGAAATTAACAAACGA 4427
Qy 4456 GGACAGCAGAGGCTGGAGACAGAGAAATGATGTTCTTGAAGGAGTCAAAAGTCTCATTTAT 4515
Db 4428 GGACAGCAGAGGCTGGAGACAGAGAAATGATGTTCTTGAAGGAGTCAAAAGTCTCATTTAT 4487

Qy 4516 GACATGTAAGCCAGAGAAAGTGGACAGAGGCCGCTGGAAGAAATTTGAGCTGTGCTTCCC 4575
Db 4488 GACATGTAAGCCAGAGAAAGTGGACAGAGGCCGCTGGAAGAAATTTGAGCTGTGCTTCCC 4547
Qy 4576 GACGGGAGTGTATCTATCTCAATGGGCGCTTGGTGTCCGAACCTCGAAATPACAGCCAA 4635
Db 4548 GACGGGAGTGTATCTATCTCAATGGGCGCTTGGTGTCCGAACCTCGAAATPACAGCCAA 4607
Qy 4636 GCAGATGTCCCATATACATCTGAAGATGAATCTCACCCGACACCACTGCTGAGCCGGG 4695
Db 4608 GCAGATGTCCCATATACATCTGAAGATGAATCTCACCCGACACCACTGCTGAGCCGGG 4667
Qy 4696 AGAACCTCTACTTGTGCTAGCTCCAGCTTCCGACAAACAGCGCTGAGTCAACGCTTTA 4755
Db 4668 AGAACCTCTACTTGTGCTAGCTCCAGCTTCCGACAAACAGCGCTGAGTCAACGCTTTA 4727
Qy 4756 GAATCAGTTTCCAGGTGGAGAGATTTCTAAGGAAAAACAGAGAGTGAATGCTAAACTG 4815
Db 4728 GAATCAGTTTCCAGGTGGAGAGATTTCTAAGGAAAAACAGAGAGTGAATGCTAAACTG 4787
Qy 4816 CTTGSAATCTCCCTGAGAACTGGAAGTGAATGACCGTCTAGACATGAATGCAACCTG 4875
Db 4788 CTTGSAATCTCCCTGAGAACTGGAAGTGAATGACCGTCTAGACATGAATGCAACCTG 4847
Qy 4876 CCGTTGAGTACAGAGGTGTGTGTGTGGGACCGAGAGAGGCTCTACCGCTGATGTC 4935
Db 4848 CCGTTGAGTACAGAGGTGTGTGTGTGGGACCGAGAGAGGCTCTACCGCTGATGTC 4907
Qy 4936 TTGAAAAATCTCCTTAACCCATGTCCAGAAATTTGAGACATCTTCCAAATTTATATATC 4995
Db 4908 TTGAAAAATCTCCTTAACCCATGTCCAGAAATTTGAGACATCTTCCAAATTTATATATC 4967
Qy 4996 AAGGACCTGAGAGAGTACTCATGATAGCAGAGAAAGCGGGCACTGTGCTTGTGAGAC 5055
Db 4968 AAGGACCTGAGAGAGTACTCATGATAGCAGAGAAAGCGGGCACTGTGCTTGTGAGAC 5027
Qy 5056 GTGAAAGAAATGAAACAGTCCCTGAGCAGTCCCACTGCTGACAGCCGACATCTCA 5115
Db 5028 GTGAAAGAAATGAAACAGTCCCTGAGCAGTCCCACTGCTGACAGCCGACATCTCA 5087
Qy 5116 CCCAATCTTTTGAAGCTGTCAAGAGGCTGCCACTTGTGTGGGGAGGCAAGATGGAAGAC 5175
Db 5088 CCCAATCTTTTGAAGCTGTCAAGAGGCTGCCACTTGTGTGGGGAGGCAAGATGGAAGAC 5147
Qy 5176 GGGCTGTGACTGTGTGAGGCAATGCCAGCAAGTGTGCTATTCCTCGCTACCAAGAAAC 5235
Db 5148 GGGCTGTGACTGTGTGAGGCAATGCCAGCAAGTGTGCTATTCCTCGCTACCAAGAAAC 5207
Qy 5236 CTGAGCAATACTGCACTCCGAGAAAGAGATAGAGACTCAGAGCCCTGAGCTGTATCCAC 5295
Db 5208 CTGAGCAATACTGCACTCCGAGAAAGAGATAGAGACTCAGAGCCCTGAGCTGTATCCAC 5267
Qy 5296 TTCAACCAATTAAGATCTCTCATTTGGAACCAATTAATTTCAACGAATGACATAGACAG 5355
Db 5268 TTCAACCAATTAAGATCTCTCATTTGGAACCAATTAATTTCAACGAATGACATAGACAG 5327
Qy 5356 TACACGCTTCAGAGAAATTCCTGATTAAGATGACCATTCCTTGGACCTGCTGTGTTGCC 5415
Db 5328 TACACGCTTCAGAGAAATTCCTGATTAAGATGACCATTCCTTGGACCTGCTGTGTTGCC 5387
Qy 5416 GCGTCTTCCAAACAGCTTCCCTGTCTCAATGCTGAGGTGAACAAGCGAGGCAAGCGAGAG 5475
Db 5388 GCGTCTTCCAAACAGCTTCCCTGTCTCAATGCTGAGGTGAACAAGCGAGGCAAGCGAGAG 5447
Qy 5476 GAGTACTTCTGTGTGTTTCAAGAAATTTGAGAGTTCGTGATTTCTTAACGAAGAGCTAGC 5535
Db 5448 GAGTACTTCTGTGTGTTTCAAGAAATTTGAGAGTTCGTGATTTCTTAACGAAGAGCTAGC 5507
Qy 5536 CGCAGACAGATCTCAAGTGAAGTGCCTTATCCTTGGCTTTGGCTTACAGAAACCCAT 5595
Db 5508 CGCAGACAGATCTCAAGTGAAGTGCCTTATCCTTGGCTTTGGCTTACAGAAACCCAT 5567

QY 5596 CTGTTGNGACCCACTTCAACTCACTGGAAGTAATTGATCCAGGACCGTCTCCAGCA 5655
| | | | |
Db 5568 CTGTTTGTACCCACTTCAACTCACTGGAAGTAATTGATCCAGGACCGTCTCCAGCA 5627
| | | | |
QY 5656 GGGACCCCTGCGGAGGTAAGTGAATCCGGAACCGCGCTACTTGGGCTGTGCATT 5715
| | | | |
Db 5628 GGGACCCCTGCGGAGGTAAGTGAATCCGGAACCGCGCTACTTGGGCTGTGCATT 5687
| | | | |
QY 5716 TCCCTCAGGAGGATTTACTTGGGCTCTCATACAGAGTAATTAAGGGTCAATTGTGTC 5775
| | | | |
Db 5688 TCCCTCAGGAGGATTTACTTGGGCTCTCATACAGAGTAATTAAGGGTCAATTGTGTC 5747
| | | | |
QY 5776 AAGGAAACCTCGTGAAGAGTCCGGACCTGAACACACGCGGCGCTGCCACTCCCGC 5835
| | | | |
Db 5748 AAGGAAACCTCGTGAAGAGTCCGGACCTGAACACACGCGGCGCTGCCACTCCCGC 5807
| | | | |
QY 5836 AGGACCCCAACAACGAGGCGCCACCAAGTACACAGACACATCAACAGCGGTGGCC 5895
| | | | |
Db 5808 AGGACCCCAACAACGAGGCGCCACCAAGTACACAGACACATCAACAGCGGTGGCC 5867
| | | | |
QY 5896 TCCAGCCCAAGCGCCCGCTGAAGCGCCAGCCACCGGAGAGCCAGACACCCACCGC 5955
| | | | |
Db 5868 TCCAGCCCAAGCGCCCGCTGAAGCGCCAGCCACCGGAGAGCCAGACACCCACCGC 5927
| | | | |
QY 5956 TACCGGAGGAGGAGCCGAGCTGCGGAGGAGCAAGTCTCGGCGCGCGCTGGAGCGA 6015
| | | | |
Db 5928 TACCGGAGGAGGAGCCGAGCTGCGGAGGAGCAAGTCTCGGCGCGCGCTGGAGCGA 5987
| | | | |
QY 6016 GAGAAGTCCCGCGCGGATCTCAGCACGCGAGAGAGCGGTCCCGCGAGAGGTCTT 6075
| | | | |
Db 5988 GAGAAGTCCCGCGCGGATCTCAGCACGCGAGAGAGCGGTCCCGCGAGAGGTCTT 6047
| | | | |
QY 6076 GAAGACAGCAGCAGGCGCGCTGCTGCGGAGCGGTAGAGACCCCGCTGTCCAGGTG 6135
| | | | |
Db 6048 GAAGACAGCAGCAGGCGCGCTGCTGCGGAGCGGTAGAGACCCCGCTGTCCAGGTG 6107
| | | | |
QY 6136 AACCAAGTCTGGGAGCCAGTCTTCAATTA 6165
| | | | |
Db 6108 AACCAAGTCTGGGAGCCAGTCTTCAATTA 6137
| | | | |
RESULT 11
AR534559
LOCUS AR534559 5877 bp DNA linear PAT 08-OCT-2004
DEFINITION Sequence 3 from patent US 6734009.
ACCESSION AR534559
VERSION AR534559.1 GI:53924886
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE
1 (bases 1 to 5877)
AUTHORS Yu,X.S., Miranda,M. and Friedle,C.J.
TITLE Human kinases and polynucleotides encoding the same
JOURNAL Patent: US 6734009-A 3 11-MAY-2004;
FEATURES
Location/Qualifiers
1..5877
/organism="Unknown"
/mol_type="genomic DNA"
ORIGIN
Query Match 94.7%; Score 5837; DB 6; Length 5877;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 5837; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGTTGAAGTTCAATATGAGCGGGAATCCCTTGGATGCTGCTGCTGAACCCATT 60
| | | | |
Db 1 ATGTTGAAGTTCAATATGAGCGGGAATCCCTTGGATGCTGCTGCTGAACCCATT 60
| | | | |
QY 61 GCCAGCCCGGCTCTCAGGCTGAATCTGTTCTTCAAGGGAACACCCCTTATGACTCA 120
| | | | |
Db 61 GCCAGCCCGGCTCTCAGGCTGAATCTGTTCTTCAAGGGAACACCCCTTATGACTCA 120
| | | | |

QY 121 CAGCAGATGTCCTCTTTTCCCGAAGGAGATTTAGATGCCCTCTTGTCTCTTTGAA 180
| | | | |
Db 121 CAGCAGATGTCCTCTTTTCCCGAAGGAGATTTAGATGCCCTCTTGTCTCTTTGAA 180
| | | | |
QY 181 GAATGAGTGAAGCTGCTGATGAAGATTAAGCAGGTGACCACTTGTCCGGAAGTAT 240
| | | | |
Db 181 GAATGAGTGAAGCTGCTGATGAAGATTAAGCAGGTGACCACTTGTCCGGAAGTAT 240
| | | | |
QY 241 TCCGACACCATAGCTAGTTACAGAGCTCCAGCTTCGCGCAAGACCTTGAAGTACA 300
| | | | |
Db 241 TCCGACACCATAGCTAGTTACAGAGCTCCAGCTTCGCGCAAGACCTTGAAGTACA 300
| | | | |
QY 301 AGCTTGTAGATTTGTGTCTCTTTGAGAGTCAAGTGTGTAAGAGCAACCGGG 360
| | | | |
Db 301 AGCTTGTAGATTTGTGTCTCTTTGAGAGTCAAGTGTGTAAGAGCAACCGGG 360
| | | | |
QY 361 GACATCTATGCTATGAAGATGAAGAAAGAGGCTTTATGGCCAGGAGGATTTCA 420
| | | | |
Db 361 GACATCTATGCTATGAAGATGAAGAAAGAGGCTTTATGGCCAGGAGGATTTCA 420
| | | | |
QY 421 TTTTGAAGAGGAGGAGCAATATTTATCTGAAGACAAAGCCGTGATCCCAATTA 480
| | | | |
Db 421 TTTTGAAGAGGAGGAGCAATATTTATCTGAAGACAAAGCCGTGATCCCAATTA 480
| | | | |
QY 481 CAGTATGCTTTTGAAGACAAAATCACTTTATCTGCTATGGAATATCAGCTGAGGG 540
| | | | |
Db 481 CAGTATGCTTTTGAAGACAAAATCACTTTATCTGCTATGGAATATCAGCTGAGGG 540
| | | | |
QY 541 GACTTCTGCTCTTTTGAATATGATGAGACCAATTTATGTAAGAAACCTGATACGTTT 600
| | | | |
Db 541 GACTTCTGCTCTTTTGAATATGATGAGACCAATTTATGTAAGAAACCTGATACGTTT 600
| | | | |
QY 601 TACCTAGCTGAGCTGATTTTGGCTTTCAAGGCTCATCTGATGGAATACGTGATCGA 660
| | | | |
Db 601 TACCTAGCTGAGCTGATTTTGGCTTTCAAGGCTCATCTGATGGAATACGTGATCGA 660
| | | | |
QY 661 GACATTAAGCTGAGAACATTTCTGTTGACCGGACAGACATCAAGCTGTGATTTT 720
| | | | |
Db 661 GACATTAAGCTGAGAACATTTCTGTTGACCGGACAGACATCAAGCTGTGATTTT 720
| | | | |
QY 721 GGAATCGCGCGGAGAAATGAATTCAAACAGATGTAATGCCAAATCCCGATGGAGAC 780
| | | | |
Db 721 GGAATCGCGCGGAGAAATGAATTCAAACAGATGTAATGCCAAATCCCGATGGAGAC 780
| | | | |
QY 781 CCAAGTTACATGAGCTCTGTAAGTGTGACTGTGATGAAGAGGAGTGAAGACCTTAC 840
| | | | |
Db 781 CCAAGTTACATGAGCTCTGTAAGTGTGACTGTGATGAAGAGGAGTGAAGACCTTAC 840
| | | | |
QY 841 GGCCTGAGCTGTGACTGTGTGCTGAGTGGGCTGATTTGCTTATGAGATGATTAAGGAGA 900
| | | | |
Db 841 GGCCTGAGCTGTGACTGTGTGCTGAGTGGGCTGATTTGCTTATGAGATGATTAAGGAGA 900
| | | | |
QY 901 TCCCGCTTCGAGAGGAAACCTCTGCGAGAACCTTCAATTAATTAAGATTTTCCAGCGG 960
| | | | |
Db 901 TCCCGCTTCGAGAGGAAACCTCTGCGAGAACCTTCAATTAATTAAGATTTTCCAGCGG 960
| | | | |
QY 961 TTTTGAATTTTCAGATGACCCCAAGAGTGAAGCTTCTTGTGATCTGATTTCAAGC 1020
| | | | |
Db 961 TTTTGAATTTTCAGATGACCCCAAGAGTGAAGCTTCTTGTGATCTGATTTCAAGC 1020
| | | | |
QY 1021 TTGTTGTGCGGCAGAAAGAGAGTGAAGTTGAAGGCTTTGTGCTGATCTGATTTCAAGC 1080
| | | | |
Db 1021 TTGTTGTGCGGCAGAAAGAGAGTGAAGTTGAAGGCTTTGTGCTGATCTGATTTCAAGC 1080
| | | | |
QY 1081 TCTAAATTTGACTGGAACAACTTGTAACTCTCTCCCTTGTGCTCAACCTCTCAAG 1140
| | | | |
Db 1081 TCTAAATTTGACTGGAACAACTTGTAACTCTCTCCCTTGTGCTCAACCTCTCAAG 1140
| | | | |
QY 1141 TCTGACATGACACCTTCAATTTTATGATGAACAGAGAGAAATTCGTGGGTTTCACTCT 1200
| | | | |
Db 1141 TCTGACATGACACCTTCAATTTTATGATGAACAGAGAGAAATTCGTGGGTTTCACTCT 1200
| | | | |
QY 1201 CCGTGCAAGCTGAGCCCTCTCAGGCTTCTCGGGTGAAGAACTGCCGTTGTGGGGTTTTCG 1260
| | | | |

|||||
Db 1201 CCGGCGAGCTGAGCCCCCTCAGGCTTCTCGGCTGAAGAACTGCGGTTGTGGGGTATTTCG 1260
|||
Qy 1261 TACAGCAAGGCACTGGGGATTCTTGTAAGTCTGAGTCTGTGTGTGGGCTGGAATCC 1320
|||
Db 1261 TAGGCAAGGCACTGGGGATTCTTGTAAGTCTGAGTCTGTGTGTGGGCTGGAATCC 1320
|||
Qy 1321 CTGTCAGAGTACTGCTCCATGGAAGAACTTCTCATCAAAGCAAGAGTCTCAAGAC 1380
|||
Db 1321 CTGTCAGAGTACTGCTCCATGGAAGAACTTCTCATCAAAGCAAGAGTCTCAAGAC 1380
|||
Qy 1381 TCTCAGAGCAAGTGTCAAGATGAGAGGAAATGACCCGGTTACATCGAGAGTGTCA 1440
|||
Db 1381 TCTCAGAGCAAGTGTCAAGATGAGAGGAAATGACCCGGTTACATCGAGAGTGTCA 1440
|||
Qy 1441 GAGGTGAGGCTGTGCTTAAGTCAAGAGAGGTGAGTGAAGGCTCTGAGACTCAGAGA 1500
|||
Db 1441 GAGGTGAGGCTGTGCTTAAGTCAAGAGAGGTGAGTGAAGGCTCTGAGACTCAGAGA 1500
|||
Qy 1501 TCCCTCTGAGACAGACCTTGTCTACCTACATCAAGAAATGCAAGTACTTAAGCGAAGT 1560
|||
Db 1501 TCCCTCTGAGACAGACCTTGTCTACCTACATCAAGAAATGCAAGTACTTAAGCGAAGT 1560
|||
Qy 1561 TTGGAGCAAGACCGGATGAGAGTGTCCAGAGAGATGACAAAGCACTGCACTTCTCAT 1620
|||
Db 1561 TTGGAGCAAGACCGGATGAGAGTGTCCAGAGAGATGACAAAGCACTGCACTTCTCAT 1620
|||
Qy 1621 GATATCAGAGACGAGCCGGAAGCTCCAGAAATCAAGAGAGGAGTACAGAGCTCAA 1680
|||
Db 1621 GATATCAGAGACGAGCCGGAAGCTCCAGAAATCAAGAGAGGAGTACAGAGCTCAA 1680
|||
Qy 1681 GTGGAGAAATGAGGTGATGATGAAATCAGTTGGAAGAGATCTTGTCTCAGCAAGAGA 1740
|||
Db 1681 GTGGAGAAATGAGGTGATGATGAAATCAGTTGGAAGAGATCTTGTCTCAGCAAGAGA 1740
|||
Qy 1741 CGGAGTATCTCTACGAATCTGAGACTGAGAGAGTCTGGGCTTGTCTGGAAGAAATTCAAG 1800
|||
Db 1741 CGGAGTATCTCTACGAATCTGAGACTGAGAGAGTCTGGGCTTGTCTGGAAGAAATTCAAG 1800
|||
Qy 1801 CGGAAAGCGACAGAAATGTCAGCATAACTGTGTAAGGCTAAGATCAAGGGAGACCTGAA 1860
|||
Db 1801 CGGAAAGCGACAGAAATGTCAGCATAACTGTGTAAGGCTAAGATCAAGGGAGACCTGAA 1860
|||
Qy 1861 GTGGGAGAAATATGGAATCTGGAAGATCAATCTGAGCAGCACTCAAAATTCAGAG 1920
|||
Db 1861 GTGGGAGAAATATGGAATCTGGAAGATCAATCTGAGCAGCACTCAAAATTCAGAG 1920
|||
Qy 1921 CTCCAGAGAAATCTGAGAAAGCTGTAAAGCCAGCACGAGGCCACCGAGCTCTCAG 1980
|||
Db 1921 CTCCAGAGAAATCTGAGAAAGCTGTAAAGCCAGCACGAGGCCACCGAGCTCTCAG 1980
|||
Qy 1981 AATATCCGCAAGGAAAGGAGCGAGCGAGAGGAGTGGAGAGCTGCGAAGCCGAG 2040
|||
Db 1981 AATATCCGCAAGGAAAGGAGCGAGCGAGAGGAGTGGAGAGCTGCGAAGCCGAG 2040
|||
Qy 2041 GATTCTTCTGAGAGCATCAGAAAGAGCTGTGAGAGCTGAGAACCGCCGCTATCTCTG 2100
|||
Db 2041 GATTCTTCTGAGAGCATCAGAAAGAGCTGTGAGAGCTGAGAACCGCCGCTATCTCTG 2100
|||
Qy 2101 GAGAAACAAGTAAAGAGACTAGAGACCAATGAGCGTGAAGAAAACAATGAGATGAC 2160
|||
Db 2101 GAGAAACAAGTAAAGAGACTAGAGACCAATGAGCGTGAAGAAAACAATGAGATGAC 2160
|||
Qy 2161 ATCCAGCAAAATATCCCAACAGATCCAGATGAGCTGATTAATTTCTGAGCTGAGAG 2220
|||
Db 2161 ATCCAGCAAAATATCCCAACAGATCCAGATGAGCTGATTAATTTCTGAGCTGAGAG 2220
|||
Qy 2221 AAACATCGGAGAGCCCAAGTCTCAGCCAGACCTAGAGTGCACTGAAACAGAAAGAG 2280
|||
Db 2221 AAACATCGGAGAGCCCAAGTCTCAGCCAGACCTAGAGTGCACTGAAACAGAAAGAG 2280
|||
Qy 2281 CAGCACTATGAGAAAGATTAAGTGTGAGCAATCAGATTAAGAAAGACCTGAGTAC 2340
|||

|||||
Db 2281 CAGCACTATGAGAAAGATTTAAAGTGTGAGCAATCAGATTAAGAAAGACCTGAGTAC 2340
|||
Qy 2341 AAGGAGCACTGAGAGAAATGATGACAGACAGAGAGAGGCCATGAGAGGGCAA 2400
|||
Db 2341 AAGGAGCACTGAGAGAAATGATGACAGACAGAGAGAGGCCATGAGAGGGCAA 2400
|||
Qy 2401 ATTCTACGCGAACAAGAGGGGATGATCAATGCTATGGAATTCAGAGATCAAGTCCCTGGA 2460
|||
Db 2401 ATTCTACGCGAACAAGAGGGGATGATCAATGCTATGGAATTCAGAGATCAAGTCCCTGGA 2460
|||
Qy 2461 CAGAGATTTGGAATCTGTCTGAAGCAATTAACCTTCAGCAAAATAGCATCTTTTACC 2520
|||
Db 2461 CAGAGATTTGGAATCTGTCTGAAGCAATTAACCTTCAGCAAAATAGCATCTTTTACC 2520
|||
Qy 2521 CAAGAGAACTGAAGGCCCAAGAGAGATGATTTCTGAATCTCAGGCAACAGAAATTTTAC 2580
|||
Db 2521 CAAGAGAACTGAAGGCCCAAGAGAGATGATTTCTGAATCTCAGGCAACAGAAATTTTAC 2580
|||
Qy 2581 CTGAGACACAGGCTGGGAAATTGAGAGGCCCAAGAACCGAAATCTGAGAGACCTGGAG 2640
|||
Db 2581 CTGAGACACAGGCTGGGAAATTGAGAGGCCCAAGAACCGAAATCTGAGAGACCTGGAG 2640
|||
Qy 2641 AAGATCAGCCACCAAGACCAAGTGAACAAGATGCGCTGGAATCTGAGACAAAGATTG 2700
|||
Db 2641 AAGATCAGCCACCAAGACCAAGTGAACAAGATGCGCTGGAATCTGAGACAAAGATTG 2700
|||
Qy 2701 CGGAGGTCAGTCTAGACACAGAGACAGAAATCTGAGTCTCAAGCCAGCTCAGCAGAG 2760
|||
Db 2701 CGGAGGTCAGTCTAGACACAGAGACAGAAATCTGAGTCTCAAGCCAGCTCAGCAGAG 2760
|||
Qy 2761 CTACAGCTTTCCTCTGACAGAGCGGAGTCAACAGTTGACAGCTCTGACAGCTGACCGGCG 2820
|||
Db 2761 CTACAGCTTTCCTCTGACAGAGCGGAGTCAACAGTTGACAGCTCTGACAGCTGACCGGCG 2820
|||
Qy 2821 GCCCTGAGAGCCAGCTTCGCAAGGCGAAGACAGAGCTGGAGAGACCAACAGAGAACT 2880
|||
Db 2821 GCCCTGAGAGCCAGCTTCGCAAGGCGAAGACAGAGCTGGAGAGACCAACAGAGAACT 2880
|||
Qy 2881 GAAGAGAGATTCAGGACCTCACGGCAATAGAGATGAATCCAGCCAAATTTGATGCT 2940
|||
Db 2881 GAAGAGAGATTCAGGACCTCACGGCAATAGAGATGAATCCAGCCAAATTTGATGCT 2940
|||
Qy 2941 CTTGTAACAGCTGTACTGTATATCAAGACCTGAGAGACGCTTAAACAGCTGACGAG 3000
|||
Db 2941 CTTGTAACAGCTGTACTGTATATCAAGACCTGAGAGACGCTTAAACAGCTGACGAG 3000
|||
Qy 3001 GACACGCTGAATCTCAACAAACAAATTTCTACTTGTCCAAACAACTCGATGAGGCTTCT 3060
|||
Db 3001 GACACGCTGAATCTCAACAAACAAATTTCTACTTGTCCAAACAACTCGATGAGGCTTCT 3060
|||
Qy 3061 GGCGCCAAGCAGAGATTTGTACAACCTGCAAGTGAAGTGAACATCTCCGCGGGAGATC 3120
|||
Db 3061 GGCGCCAAGCAGAGATTTGTACAACCTGCAAGTGAAGTGAACATCTCCGCGGGAGATC 3120
|||
Qy 3121 ACGGAAACGAGATGACGCTTACCAAGCAGAAAGAAAGATGAGGCTCTGAAGACCAAG 3180
|||
Db 3121 ACGGAAACGAGATGACGCTTACCAAGCAGAAAGAAAGATGAGGCTCTGAAGACCAAG 3180
|||
Qy 3181 TGCAACATGCTGAGAGAAACAGTCAATGGAATTTGAGAGCCCTTAAACGATGAGCTGAGAA 3240
|||
Db 3181 TGCAACATGCTGAGAGAAACAGTCAATGGAATTTGAGAGCCCTTAAACGATGAGCTGAGAA 3240
|||
Qy 3241 AAAGAGCGGAGTGGAGAGGCTGAGAGAGCGTCTGGGTGATGAGAAATCCCAATTTGAG 3300
|||
Db 3241 AAAGAGCGGAGTGGAGAGGCTGAGAGAGCGTCTGGGTGATGAGAAATCCCAATTTGAG 3300
|||
Qy 3301 TGTGCGGTTTGAAGAGCTGCAAGAGATGCTGGAACCGAGAAACAGACAGGCGAGAGCC 3360
|||
Db 3301 TGTGCGGTTTGAAGAGCTGCAAGAGATGCTGGAACCGAGAAACAGACAGGCGAGAGCC 3360
|||
Qy 3361 GATCAGCGGATCACCGAGTCTCCGCAAGTGTGAGAGCTGACAGTGAAGAGCAAGGCT 3420
|||
Db 3361 GATCAGCGGATCACCGAGTCTCCGCAAGTGTGAGAGCTGACAGTGAAGAGCAAGGCT 3420
|||

QY	3421	GAAATTCGTGCTTGCAGCAGGCTCTCAAAGACGAGCTGAAGGCGAGACCTCTT	3480
Db	3421	GAAATTCGTGCTTGCAGCAGGCTCTCAAAGACGAGACTGAAGGCGAGACCTCTT	3480
QY	3481	GAAACCTCAATGACCTGGAGAAAGAAAGCATGCTTAAGCTTGAATGAATGCCGAACTTA	3540
Db	3481	GAAACCTCAATGACCTGGAGAAAGAAAGCATGCTTAAGCTTGAATGAATGCCGAACTTA	3540
QY	3541	CACGAGAAGCTGAGACTGAACGAGAGCTCAACAGAGCTTCTGGAAGAGCAAGCCAA	3600
Db	3541	CACGAGAAGCTGAGACTGAACGAGAGCTCAACAGAGCTTCTGGAAGAGCAAGCCAA	3600
QY	3601	TTACACACACATGAGACCTTCGAGAAAAATCACTTTTCGTCTGACTCAAGACTGCA	3660
Db	3601	TTACACACACATGAGACCTTCGAGAAAAATCACTTTTCGTCTGACTCAAGACTGCA	3660
QY	3661	GAACTCTGATCGGGCTGATCTACTGAAGACAGAAAGAGTGACTTGGATACAGCTG	3720
Db	3661	GAACTCTGATCGGGCTGATCTACTGAAGACAGAAAGAGTGACTTGGATACAGCTG	3720
QY	3721	GAAAAATTCAAGTTCTCTAATTCATGAAAAGGTAAAAATGAAAGCACTAATTTCTCA	3780
Db	3721	GAAAAATTCAAGTTCTCTAATTCATGAAAAGGTAAAAATGAAAGCACTAATTTCTCA	3780
QY	3781	CAAAACCAACTCATTAATTTTTCGAACCCAAATGACCAACTGCTTAAAAAAGAAAG	3840
Db	3781	CAAAACCAACTCATTAATTTTTCGAACCCAAATGACCAACTGCTTAAAAAAGAAAG	3840
QY	3841	GTTCCCTGCACTAGTACATGAGCTGAAGCTGAGCCCTGAGAGAGAGAAAGCTCGCTGCA	3900
Db	3841	GTTCCCTGCACTAGTACATGAGCTGAAGCTGAGCCCTGAGAGAGAGAAAGCTCGCTGCA	3900
QY	3901	GAGCTAAGAGAAAGCCCTTTCAGAAAGCCGCATCGAGCTCCGCTCCCGAGAGAAAGCT	3960
Db	3901	GAGCTAAGAGAAAGCCCTTTCAGAAAGCCGCATCGAGCTCCGCTCCCGAGAGAAAGCT	3960
QY	3961	GCCCAACCCCAAGCAACGAGACCAACCAACCATCAACGACGACACCGAGAGACAG	4020
Db	3961	GCCCAACCCCAAGCAACGAGACCAACCAACCATCAACGACGACACCGAGAGACAG	4020
QY	4021	ATGCGCATGTCCGCATCGTGTGCGGTCCGACAGACCAAGCCCACTGACCTGACTGCTG	4080
Db	4021	ATGCGCATGTCCGCATCGTGTGCGGTCCGACAGACCAAGCCCACTGACCTGACTGCTG	4080
QY	4081	GCCCCGCATCCAGCCCGAGAAAGAGTCTTCAACTCAAGAGAAATTTAGTCGGCTCTT	4140
Db	4081	GCCCCGCATCCAGCCCGAGAAAGAGTCTTCAACTCAAGAGAAATTTAGTCGGCTCTT	4140
QY	4141	AAGAAACCATGACACCAATATTCCTCAACCGATTCAACGTAAGAGACTGAACTAGCAGCC	4200
Db	4141	AAGAAACCATGACACCAATATTCCTCAACCGATTCAACGTAAGAGACTGAACTAGCAGCC	4200
QY	4201	ACAAAGTGTGTGTGTCTGGAATACCGTGCACTTTGAGCGCCAGGCATCAATGTCTC	4260
Db	4201	ACAAAGTGTGTGTGTCTGGAATACCGTGCACTTTGAGCGCCAGGCATCAATGTCTC	4260
QY	4261	GAAATGTACGATGTGTCAACCCCAATGTCTCCACGTCTTGCAGGCCACTTGCGGCTTG	4320
Db	4261	GAAATGTACGATGTGTCAACCCCAATGTCTCCACGTCTTGCAGGCCACTTGCGGCTTG	4320
QY	4321	CCTGCTGAATATGCAACAATTACGAGAGGCTTTCGCGGTGACAAATGAATCTCCCA	4380
Db	4321	CCTGCTGAATATGCAACAATTACGAGAGGCTTTCGCGGTGACAAATGAATCTCCCA	4380
QY	4381	GGTCTCCAGACCAAGAGCCCAAGCAGTTCGACCTGGAAGGGTGAATGAAGTGTGCC	4440
Db	4381	GGTCTCCAGACCAAGAGCCCAAGCAGCAGTTCGACCTGGAAGGGTGAATGAAGTGTGCC	4440
QY	4441	AGGAATTAACAAGAGACAGCAAGGCTGGGACAGAAAGTACATTTGCTCGAGAGGATCA	4500
Db	4441	AGGAATTAACAAGAGACAGCAAGGCTGGGACAGAAAGTACATTTGCTCGAGAGGATCA	4500

QY	4501	AAAAGTCCTCATTTTATGACAAATGAAAGCCGAGAAAGCTGGACAGAGGCCGCTGTGAAGAAATTT	4560
Db	4501	AAAAGTCCTCATTTTATGACAAATGAAAGCCGAGAAAGCTGGACAGAGGCCGCTGTGAAGAAATTT	4560
QY	4561	GAGCTGTGCTTCCCGACGGGGATGTATCTATTCAATGGTGGCGTTTGGTCTTCGGAATC	4620
Db	4561	GAGCTGTGCTTCCCGACGGGGATGTATCTATTCAATGGTGGCGTTTGGTCTTCGGAATC	4620
QY	4621	GCAAAATACAGCCAAAGACAGATGTGCCATATACATGAATGAATGAAATCTCACCCGCAACC	4680
Db	4621	GCAAAATACAGCCAAAGACAGATGTGCCATATACATGAATGAATGAAATCTCACCCGCAACC	4680
QY	4681	ACCTGCTGGCCCGGGAGAAACCTCTACTTCTAGCTCCAGCTTCCCTGACAAACAGGC	4740
Db	4681	ACCTGCTGGCCCGGGAGAAACCTCTACTTCTAGCTCCAGCTTCCCTGACAAACAGGC	4740
QY	4741	TGGGTACCCGCTTAGAATCAGTTGTGGCAGATGGGAGAGTTTCTAGGAAAAAGCAGAA	4800
Db	4741	TGGGTACCCGCTTAGAATCAGTTGTGGCAGATGGGAGAGTTTCTAGGAAAAAGCAGAA	4800
QY	4801	GCTGATGCTAAACTGCTTGGAAACTCCCTGCTGAACTGGAAGGTGATGACCCGCTAGAC	4860
Db	4801	GCTGATGCTAAACTGCTTGGAAACTCCCTGCTGAAACTGGAAGGTGATGACCCGCTAGAC	4860
QY	4861	ATGAACTGCAACGCTGCCCTTCAGTGACCAAGTGTGTTGTGGGACCGAGAAAGGCTC	4920
Db	4861	ATGAACTGCAACGCTGCCCTTCAGTGACCAAGTGTGTTGTGGGACCGAGAAAGGCTC	4920
QY	4921	TACGCCCTGAAATGTCTTGAAAAACTCCCTTAACCCATGTCCAGAAATTTGAGCAGCTTTC	4980
Db	4921	TACGCCCTGAAATGTCTTGAAAAACTCCCTTAACCCATGTCCAGAAATTTGAGCAGCTTTC	4980
QY	4981	CAAAATTTATATTCACAGACCTGGAAGAACTACTCATGATAGCAGAGAAAGACGGGCA	5040
Db	4981	CAAAATTTATATTCACAGACCTGGAAGAACTACTCATGATAGCAGAGAAAGACGGGCA	5040
QY	5041	CTGTGCTTGTGGACGTGAAGAAAGTGAACAGTCCCTGGCCCAAGTCCACCTGCTGCC	5100
Db	5041	CTGTGCTTGTGGACGTGAAGAAAGTGAACAGTCCCTGGCCCAAGTCCACCTGCTGCC	5100
QY	5101	CAGCCCGACATCTACCCCAATTTTGAAGCTGTACAAAGGCGGCCACTGTTTGGGGCA	5160
Db	5101	CAGCCCGACATCTACCCCAATTTTGAAGCTGTACAAAGGCGGCCACTGTTTGGGGCA	5160
QY	5161	GGCAAGATTGGAAGACGGGCTCTGCATCTGTGACGCAATGCCCAAGCAATGCTCATTTCTC	5220
Db	5161	GGCAAGATTGGAAGACGGGCTCTGCATCTGTGACGCAATGCCCAAGCAATGCTCATTTCTC	5220
QY	5221	CGTACACACGAAAACTCAGCAAAATACCTGCATCCGGAAGAAGATAGACCTCAGAGCCC	5280
Db	5221	CGTACACACGAAAACTCAGCAAAATACCTGCATCCGGAAGAAGATAGACCTCAGAGCCC	5280
QY	5281	TGCAGCTGTATCCACTTCAACCAATTAAGTATCTCATATTGGAACCAATTAATCTTAGCAA	5340
Db	5281	TGCAGCTGTATCCACTTCAACCAATTAAGTATCTCATATTGGAACCAATTAATCTTAGCAA	5340
QY	5341	ATGCACATGAAGCAGTACAGCTCCGAGGAATTCCTGATTAAGAAACCAATTCCTTTGGCA	5400
Db	5341	ATGCACATGAAGCAGTACAGCTCCGAGGAATTCCTGATTAAGAAACCAATTCCTTTGGCA	5400
QY	5401	CCTGCTGTGTTTGCCTTCCTTCCAAACAGCTTCCCTGTCTCAATCTGTGACGGTGAACAGC	5460
Db	5401	CCTGCTGTGTTTGCCTTCCTTCCAAACAGCTTCCCTGTCTCAATCTGTGACGGTGAACAGC	5460
QY	5461	GCAAGGACAGCAAGAGAGTACTTGTCTGTGTTTCAAGAAATTTGGAGTGTTCCTGGAATCT	5520
Db	5461	GCAAGGACAGCAAGAGAGTACTTGTCTGTGTTTCAAGAAATTTGGAGTGTTCCTGGAATCT	5520
QY	5521	TACGGAAGAAGTACGACCGCAACAGCAATCAAGTGAAGTCCCTTACCTTTGGCCTTTGCC	5580
Db	5521	TACGGAAGAAGTACGACCGCAACAGCAATCAAGTGAAGTCCCTTACCTTTGGCCTTTGCC	5580
QY	5581	TACAGGAACCCCTATCTGTTTGTGACCACTTCACTCACTGAAATTAATTAGATCCAG	5640

Db 5581 TACAGAGAACCCCTATCTGTTTGATCCACTTCACTCAAGTAATTGATCCAG 5640
Qy 5641 GCAAGCTCTCAGAGAGGACCCCTGCGGAGCGTACCTGACATCCGAAACCGCGCTAC 5700
Db 5641 GCAAGCTCTCAGAGAGGACCCCTGCGGAGCGTACCTGACATCCGAAACCGCGCTAC 5700
Qy 5701 CTGGGCCCCTGCCATTTCTCAGAGAGCATTTACTTGGCGTCTCTCAATCAAGATAAATTA 5760
Db 5701 CTGGGCCCCTGCCATTTCTCAGAGAGCATTTACTTGGCGTCTCTCAATCAAGATAAATTA 5760
Qy 5761 AGGCTCTTTGCTGCAAGGAAACCTGTGAAGAGTCCGCACTGAACCAACCGGGGC 5820
Db 5761 AGGCTCTTTGCTGCAAGGAAACCTGTGAAGAGTCCGCACTGAACCAACCGGGGC 5820
Qy 5821 CCGTCCACTCCGCGAG 5837
Db 5821 CCGTCCACTCCGCGAG 5837

RESULT 12

AX574427 5877 bp DNA linear PAT 07-JAN-2003
LOCUS AX574427
DEFINITION Sequence 3 from Patent W002059325.
ACCESSION AX574427
VERSION AX574427.1 GI:27551752
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Homo sapiens
Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
AUTHORS Yu,X., Miranda,M. and Friddele,C.J.
TITLE Human kinases and polynucleotides encoding the same
JOURNAL Patent: WO 02059325-A 3 01-AUG-2002;
Lexicon Genetics Incorporated (US)
FEATURES
Location/Qualifiers
source
1..5877
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Query Match 94.7%; Score 5837; DB 6; Length 5877;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 5837; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGTTGAAGTTCAAAATATGAGCGCGAATCTTTGATGCTGTGCTGAAACCAAT 60
Db 1 ATGTTGAAGTTCAAAATATGAGCGCGAATCTTTGATGCTGTGCTGAAACCAAT 60
Qy 61 GCCAGCGCGGCTCCAGGCTGAATCTGTTCTTCCAGGAGAAACCAACCTTATGACTCAA 120
Db 61 GCCAGCGCGGCTCCAGGCTGAATCTGTTCTTCCAGGAGAAACCAACCTTATGACTCAA 120
Qy 121 CAGCAGATGCTCTCTCTTCCCGAGAGGATATTAGATGCCCTTTGTTCTTTGAA 180
Db 121 CAGCAGATGCTCTCTCTTCCCGAGAGGATATTAGATGCCCTTTGTTCTTTGAA 180
Qy 181 GAATGAGTCAAGCTGCTGTGATGAAGATTAAAGCAGTGAACAATTGTCCGGAATAT 240
Db 181 GAATGAGTCAAGCTGCTGTGATGAAGATTAAAGCAGTGAACAATTGTCCGGAATAT 240
Qy 241 TCCGACACCATAGTGAAGTTAAAGAGACTCCAGCTTCCGCAAGAGCTTGAAGTACA 300
Db 241 TCCGACACCATAGTGAAGTTAAAGAGACTCCAGCTTCCGCAAGAGCTTGAAGTACA 300
Qy 301 AGCTTTAGAGTTGTGCTCACTTGTCTGAAGTGAAGTGAAGAGAAAGCAACCGGG 360
Db 301 AGCTTTAGAGTTGTGCTCACTTGTCTGAAGTGAAGTGAAGAGAAAGCAACCGGG 360
Qy 361 GACATCTATGCTATGAAGATGAAGAAAGGCTTTATTTGGCCAGAGAGCTTTCA 420
Db 361 GACATCTATGCTATGAAGATGAAGAAAGGCTTTATTTGGCCAGAGAGCTTTCA 420

Db 361 GACATCTATGCTATGAAGATGAAGAAAGGCTTTATTTGGCCAGAGAGCTTTCA 420
Qy 421 TTTTGGAGAGAGCGGAA CATATTATCTCGAAGCAAGACCGTGAAATCCCAATTA 480
Db 421 TTTTGGAGAGAGCGGAA CATATTATCTCGAAGCAAGACCGTGAAATCCCAATTA 480
Qy 481 CAGTATGCCCTTTCAAGACAAAATTCACCTTTATCTGGTCAATGAAATTCAGCCGTGAGGG 540
Db 481 CAGTATGCCCTTTCAAGACAAAATTCACCTTTATCTGGTCAATGAAATTCAGCCGTGAGGG 540
Qy 541 GACTTGCTGACCTTTTGAATAGATATGAGAACCAAGTATAGTAAACCTGATCAGTTT 600
Db 541 GACTTGCTGACCTTTTGAATAGATATGAGAACCAAGTATAGTAAACCTGATCAGTTT 600
Qy 601 TACCTAGCTGAGCTGATTTTGGCTGTTCAAGGCTTCATCTGATGAGATACGTGCATCGA 660
Db 601 TACCTAGCTGAGCTGATTTTGGCTGTTCAAGGCTTCATCTGATGAGATACGTGCATCGA 660
Qy 661 GACATCAAGCTGAGAACATTCTGTTGACCGCAAGAGCAACATCAAGCTGTGAGATTTT 720
Db 661 GACATCAAGCTGAGAACATTCTGTTGACCGCAAGAGCAACATCAAGCTGTGAGATTTT 720
Qy 721 GATCTGCGCGGAAATGAATTCAAACAGATGATGCAAACTCCGATTTGGAGC 780
Db 721 GATCTGCGCGGAAATGAATTCAAACAGATGATGCAAACTCCGATTTGGAGC 780
Qy 781 CCAGATTTACATGAGCTCTCTGATGAGTCTGATGATGAACCGGAGATGAAAGGACCTAC 840
Db 781 CCAGATTTACATGAGCTCTCTGATGAGTCTGATGATGAACCGGAGATGAAAGGACCTAC 840
Qy 841 GGCCTGAGCTGATGATGCTGTGATGAGTGGCGGTGATTCCTATGATGATTTATGGAGA 900
Db 841 GGCCTGAGCTGATGATGCTGTGATGAGTGGCGGTGATTCCTATGATGATTTATGGAGA 900
Qy 901 TCCCGCTTCCAGAGGGAACCTGCGCAGAACCTTCAATTAATTGAATTTCCAGCGG 960
Db 901 TCCCGCTTCCAGAGGGAACCTGCGCAGAACCTTCAATTAATTGAATTTCCAGCGG 960
Qy 961 TTTTGAATTTCCAGATGACCCCAAGTGAAGTGAAGTCTTCTGATCTGATTTCAAAAGC 1020
Db 961 TTTTGAATTTCCAGATGACCCCAAGTGAAGTGAAGTCTTCTGATCTGATTTCAAAAGC 1020
Qy 1021 TTGTTGCGGCGCAAGAGAGAGAGCTGAAGTGAAGTCTTCTGATCTGATTTCTTTC 1080
Db 1021 TTGTTGCGGCGCAAGAGAGAGAGCTGAAGTGAAGTCTTCTGATCTGATTTCTTTC 1080
Qy 1081 TCTTAAATTAATCTGGAACCAATTCGTAATCTCTCCCGCTTCTGATCTGATTTCTTTC 1140
Db 1081 TCTTAAATTAATCTGGAACCAATTCGTAATCTCTCCCGCTTCTGATCTGATTTCTTTC 1140
Qy 1141 TCTGACATGACACTTCAATTTTGAATGAACAGAGAAATTCGTTGATTTCACTCTCT 1200
Db 1141 TCTGACATGACACTTCAATTTTGAATGAACAGAGAAATTCGTTGATTTCACTCTCT 1200
Qy 1201 CCGTGCAGCTGAGCCCTCAGGCTTCTCGGATGAAGAACTGCGTTGTGGGGTTTTTCG 1260
Db 1201 CCGTGCAGCTGAGCCCTCAGGCTTCTCGGATGAAGAACTGCGTTGTGGGGTTTTTCG 1260
Qy 1261 TACAGCAAGGCACTGGGAGTTCTTGTGATATCTGAGTCTGTGTGTGGGTCTGAGATCC 1320
Db 1261 TACAGCAAGGCACTGGGAGTTCTTGTGATATCTGAGTCTGTGTGTGGGTCTGAGATCC 1320
Qy 1321 CCGTGCAGGCTGAGCTCAGTGAAGAACTTCTCATCAAAAGCAAGAGCTCAAGAC 1380
Db 1321 CCGTGCAGGCTGAGCTCAGTGAAGAACTTCTCATCAAAAGCAAGAGCTCAAGAC 1380
Qy 1381 TCTCAGAGCAAGTGTCACAAGATGAGCAAGAAATGACCCGGTTTACATGAGAGTGTCA 1440
Db 1381 TCTCAGAGCAAGTGTCACAAGATGAGCAAGAAATGACCCGGTTTACATGAGAGTGTCA 1440
Qy 1441 GAGTGAAGGCTGTCTTATCTAGAGAGAGTGAAGCTGAAAGGCTTGAAGCTCAAGAGA 1500
Db 1441 GAGTGAAGGCTGTCTTATCTAGAGAGAGTGAAGCTGAAAGGCTTGAAGCTCAAGAGA 1500

QY 1501 TCCTCCCTGAGCAGACCTTGTACCTACATCAACAAATGCAATGCTTAAAGCAAGT 1560
DB 1501 TCCTCCCTGAGCAGACCTTGTACCTACATCAACAAATGCAATGCTTAAAGCAAGT 1560
QY 1561 TTGAGCAGACGACGATGAGAGGTGTCCAGAGAGATGACAAAGACCTGCAGCTTCTCAT 1620
DB 1561 TTGAGCAGACGACGATGAGAGGTGTCCAGAGAGATGACAAAGACCTGCAGCTTCTCAT 1620
QY 1621 GATTCAGAGACGACGACCGGAGCTCCAAAGAAATCAAAAGACGAGATCCAGGCTCAA 1680
DB 1621 GATTCAGAGACGACGACCGGAGCTCCAAAGAAATCAAAAGACGAGATCCAGGCTCAA 1680
QY 1681 GTGGAAGAAATGAGGTGATGATGAATCAAGTTGAAAGAGATCTTGTCTCAGCAAGAGA 1740
DB 1681 GTGGAAGAAATGAGGTGATGATGAATCAAGTTGAAAGAGATCTTGTCTCAGCAAGAGA 1740
QY 1741 CGGAGTATCTTACGATCTGAGCTGAGAGAGTCTGGCTTGTGCTGAGAAATTCAG 1800
DB 1741 CGGAGTATCTTACGATCTGAGCTGAGAGAGTCTGGCTTGTGCTGAGAAATTCAG 1800
QY 1801 CGGAAACGACAGAAATGTCAGCATTAACCTGTTGAAGCTTAAGATCAAGGGAAGCTGAA 1860
DB 1801 CGGAAACGACAGAAATGTCAGCATTAACCTGTTGAAGCTTAAGATCAAGGGAAGCTGAA 1860
QY 1861 GTGGAAGAAATGAGGTGATGATGAATCAAGTTGAAAGAGATCTTGTCTCAGCAAGAGA 1920
DB 1861 GTGGAAGAAATGAGGTGATGATGAATCAAGTTGAAAGAGATCTTGTCTCAGCAAGAGA 1920
QY 1921 CTCGAAGAGAACTGAGAGAGGCTGTAAAGCCAGACCGAGGCTGAGGCTGCTGAG 1980
DB 1921 CTCGAAGAGAACTGAGAGAGGCTGTAAAGCCAGACCGAGGCTGAGGCTGCTGAG 1980
QY 1981 AATATCCGCGCAGGCAAGAGACGAGCCGAGAGAGAGCTGAGAGAGCTGCAAGACCGAGAG 2040
DB 1981 AATATCCGCGCAGGCAAGAGACGAGCCGAGAGAGAGCTGAGAGAGCTGCAAGACCGAGAG 2040
QY 2041 GATTCCTTCTGAGAGCATCAGAAAGAGCTGTGAGAGCTGAGAGAGCTGAGAGAGCTGAG 2100
DB 2041 GATTCCTTCTGAGAGCATCAGAAAGAGCTGTGAGAGCTGAGAGAGCTGAGAGAGCTGAG 2100
QY 2101 GAGAACAGAGTAAAGAGACTGAGAGACCATGAGAGCTGAGAGAGAGCTGAGAGAGCTGAG 2160
DB 2101 GAGAACAGAGTAAAGAGACTGAGAGACCATGAGAGCTGAGAGAGAGCTGAGAGAGCTGAG 2160
QY 2161 ATCCAGACAAATCCCAACAGATCCAGAGAGCTGAGAGAGCTGAGAGAGCTGAGAGAG 2220
DB 2161 ATCCAGACAAATCCCAACAGATCCAGAGAGCTGAGAGAGCTGAGAGAGCTGAGAGAG 2220
QY 2221 AAAACATCGGAGGCGCAAGTCTCAAGCCAGACCTTGAAGTGCACCTGAAACAGAAAGAG 2280
DB 2221 AAAACATCGGAGGCGCAAGTCTCAAGCCAGACCTTGAAGTGCACCTGAAACAGAAAGAG 2280
QY 2281 CAGCACTATGAGAGAAAGATTAAAGTGTGACAAATCAGTAAAGAAAGACCTGAGTAC 2340
DB 2281 CAGCACTATGAGAGAAAGATTAAAGTGTGACAAATCAGTAAAGAAAGACCTGAGTAC 2340
QY 2341 AAGAGACACTGAGAGACATGATGACAGACAGAGAGAGGCGCATGAGAGAGGCAAA 2400
DB 2341 AAGAGACACTGAGAGACATGATGACAGACAGAGAGAGGCGCATGAGAGAGGCAAA 2400
QY 2401 ATTCCTCAGGAGCAGAGGCGATGATCAATGCTATGATTCCAAGATCCAGATCCCTGAA 2460
DB 2401 ATTCCTCAGGAGCAGAGGCGATGATCAATGCTATGATTCCAAGATCCAGATCCCTGAA 2460
QY 2461 CAGAGGATGAGAGACTGTCTGAGAGCAATTAACCTGAGCAAAATGAGCTCTTTTAC 2520
DB 2461 CAGAGGATGAGAGACTGTCTGAGAGCAATTAACCTGAGCAAAATGAGCTCTTTTAC 2520
QY 2521 CAAAGGAAATGAGGCGCAAGAGAGATGATTTCTGAATCAGGCAAGAGAAATTTTAC 2580
DB 2521 CAAAGGAAATGAGGCGCAAGAGAGATGATTTCTGAATCAGGCAAGAGAAATTTTAC 2580

QY 2581 CTGAGACACAGGCTGGGAAAGTTGAGAGCCAGAACCCGAAACTGAGAGACAGCTGAG 2640
DB 2581 CTGAGACACAGGCTGGGAAAGTTGAGAGCCAGAACCCGAAACTGAGAGACAGCTGAG 2640
QY 2641 AAGATCAGCCACCAAGACCAAGTGAACAAGATGAGGTGCTGAGAACTGAGAGCAAGATTG 2700
DB 2641 AAGATCAGCCACCAAGACCAAGTGAACAAGATGAGGTGCTGAGAACTGAGAGCAAGATTG 2700
QY 2701 CGGAGGTCAGTCTAGAGCAGAGAGCAGAGAACTGAGAGCTCAAGCCAGCTCAAGAG 2760
DB 2701 CGGAGGTCAGTCTAGAGCAGAGAGCAGAGAACTGAGAGCTCAAGCCAGCTCAAGAG 2760
QY 2761 CTACAGCTTCTCCCTGAGAGAGCCGAGTCAAGTTGAACAGCCTGACAGGCTGACGGGCG 2820
DB 2761 CTACAGCTTCTCCCTGAGAGAGCCGAGTCAAGTTGAACAGCCTGACAGGCTGACGGGCG 2820
QY 2821 GCCCTGAGAGCCAGCTTGGCCAGGGGAAACAGAGCTGAGAGAGCAAGAGAGCT 2880
DB 2821 GCCCTGAGAGCCAGCTTGGCCAGGGGAAACAGAGCTGAGAGAGCAAGAGAGCT 2880
QY 2881 GAAAGAGATCCAGGACCTCAGGCAATAGAGATGAAATCCAGGCAAAATTTGATGCT 2940
DB 2881 GAAAGAGATCCAGGACCTCAGGCAATAGAGATGAAATCCAGGCAAAATTTGATGCT 2940
QY 2941 CTTCTGTAACAGCTGTACTGTATCAAGACCTGAGAGAGAGCTTAAACAGCTGACGAG 3000
DB 2941 CTTCTGTAACAGCTGTACTGTATCAAGACCTGAGAGAGAGCTTAAACAGCTGACGAG 3000
QY 3001 GACAGGCTGAACTCAACCAACCAAACTTCTACTGTCCAAACAACTCCGATGAGGCTTCT 3060
DB 3001 GACAGGCTGAACTCAACCAACCAAACTTCTACTGTCCAAACAACTCCGATGAGGCTTCT 3060
QY 3061 GGGCCACACAGCAGATTGTACAACTGCGAAGTGAAGTGAAGCCATCTCCGCGGAGATC 3120
DB 3061 GGGCCACACAGCAGATTGTACAACTGCGAAGTGAAGTGAAGCCATCTCCGCGGAGATC 3120
QY 3121 ACCGAAACGAGATGACGCTTACACAGCCAGAGCAAAACGATGAGGCTGTAAAGCACG 3180
DB 3121 ACCGAAACGAGATGACGCTTACACAGCCAGAGCAAAACGATGAGGCTGTAAAGCACG 3180
QY 3181 TGCAACATCTGAGAGAACAGGTCAATGATTTGAGAGCCCTTAAACGATGAGCTGTAGAA 3240
DB 3181 TGCAACATCTGAGAGAACAGGTCAATGATTTGAGAGCCCTTAAACGATGAGCTGTAGAA 3240
QY 3241 AAAAGCCGACAGTGGAGGCTGAGAGAGAGCTCTGGGTGATGAGAAATCCAGTTTGA 3300
DB 3241 AAAAGCCGACAGTGGAGGCTGAGAGAGAGCTCTGGGTGATGAGAAATCCAGTTTGA 3300
QY 3301 TGTGGGTTCGAGAGCTGAGAGATGCTGAGACACCGAAGAACAGACAGGCGAGAGCC 3360
DB 3301 TGTGGGTTCGAGAGCTGAGAGATGCTGAGACACCGAAGAACAGACAGGCGAGAGCC 3360
QY 3361 GATCAGCGGATCAACGAGCTGCGCAGAGGTGAGAGCTGAGAGAGAGCAAGAGCT 3420
DB 3361 GATCAGCGGATCAACGAGCTGCGCAGAGGTGAGAGCTGAGAGAGAGCAAGAGCT 3420
QY 3421 GAGATTCGCTCTGAGAGAGCTCTCAAGAGCAGAACTGAAAGCCAGAGCTCTCT 3480
DB 3421 GAGATTCGCTCTGAGAGAGCTCTCAAGAGCAGAACTGAAAGCCAGAGCTCTCT 3480
QY 3481 GACAACTCAATGACCTGAGAGAGAGCATGCTTATGATGAATGATGATCCGAGACTTA 3540
DB 3481 GACAACTCAATGACCTGAGAGAGAGCATGCTTATGATGAATGATGATCCGAGACTTA 3540
QY 3541 CAGCAGAGCTGAGAGCTGAGAGAGAGCTCAAAACAGAGGCTTCTGAGAGAGCAAGCTCAA 3600
DB 3541 CAGCAGAGCTGAGAGCTGAGAGAGAGCTCAAAACAGAGGCTTCTGAGAGAGCAAGCTCAA 3600
QY 3601 TTACAGCAGCAGATGAGCCTGAGAGAAATTCATTTTCGCTGACTCAAGGCTGCA 3660
DB 3601 TTACAGCAGCAGATGAGCCTGAGAGAAATTCATTTTCGCTGACTCAAGGCTGCA 3660
QY 3661 GAAAGCTTAGATCGGCTGATCTACTGAAAGCAGAAAGATGACTTGAATGATCAGCTG 3720

Db 3661 GAAGCTCTGATGGGCGTGTACTAGAGACAGAAAGAGTGTGAGTACAGCTG 3720
Qy 3721 GAAACATTGAGTTCTCTATTCATGAAAGTGAAGTGAAGGCACTATTTCCAA 3780
Db 3721 GAAAACTTCAGTTCTCTATTTCTCATGAAAAGGTGAAGTGAAGGCACTATTTCTCAA 3780
Qy 3781 CAAAACCAATCTATGATTTTCTGCAAGCCAAAATGACCAACCTGTGTAAGAAAAAG 3840
Db 3781 CAAAACCAATCTATGATTTTCTGCAAGCCAAAATGACCAACCTGTGTAAGAAAAAG 3840
Qy 3841 GTTCTCTGAGTACATGAGCTGAACTGGCTTGAAGAAAGAAAGTCCGTGCA 3900
Db 3841 GTTCTCTGAGTACATGAGCTGAACTGGCTTGAAGAAAGAAAGTCCGTGCA 3900
Qy 3901 GAGCTAGAGAGAGCCCTTCAGAAAGACCGGATCGAGCTCGGTCGCGCCGAGAGAACT 3960
Db 3901 GAGCTAGAGAGAGCCCTTCAGAAAGACCGGATCGAGCTCGGTCGCGCCGAGAGAACT 3960
Qy 3961 GCCCACCAGAAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 4020
Db 3961 GCCCACCAGAAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 4020
Qy 4021 ATGGCCATGTCGCGCATGTCGCGTCCGAGAGACCAAGCCAGTGCATGAGCTGCTG 4080
Db 4021 ATGGCCATGTCGCGCATGTCGCGTCCGAGAGACCAAGCCAGTGCATGAGCTGCTG 4080
Qy 4081 GCGCCGCGCATTCAGCCGAGAAAGAGTCTTCAACTCAGAGAAATTAAGTCGCGCTT 4140
Db 4081 GCGCCGCGCATTCAGCCGAGAAAGAGTCTTCAACTCAGAGAAATTAAGTCGCGCTT 4140
Qy 4141 AAGGAAAGCATGACACCAATATTTCTCAGCGATTCACGTAGAGCTGAACATGCGAGCC 4200
Db 4141 AAGGAAAGCATGACACCAATATTTCTCAGCGATTCACGTAGAGCTGAACATGCGAGCC 4200
Qy 4201 ACAAGTGTCTGTGTCTGATACCTGTGACCTTTGAGCGCCAGGATCAAAATGTCTC 4260
Db 4201 ACAAGTGTCTGTGTCTGATACCTGTGACCTTTGAGCGCCAGGATCAAAATGTCTC 4260
Qy 4261 GAATGTCAAGTATGTCTCAAGCCCAAGTGTCCAGCTTCCAGCACTGCGGCTTG 4320
Db 4261 GAATGTCAAGTATGTCTCAAGCCCAAGTGTCTCCAGCTTCCAGCACTGCGGCTTG 4320
Qy 4321 CTTGCTGATATATGCAACAATTCAACGAGGCTTCTGCGGTGCAAAATGAACTGCCA 4380
Db 4321 CTTGCTGATATATGCAACAATTCAACGAGGCTTCTGCGGTGCAAAATGAACTGCCA 4380
Qy 4381 GGTCTCCAGACCAAGGAGCCAGGAGGCTTGCACCTGGAAGGATGAAAGTGTCC 4440
Db 4381 GGTCTCCAGACCAAGGAGCCAGGAGGCTTGCACCTGGAAGGATGAAAGTGTCC 4440
Qy 4441 AAGGAAATCAAAAGAGACAGCAAGGCTGAGCAAGAAATGATTTCTGAGAGATCA 4500
Db 4441 AAGGAAATCAAAAGAGACAGCAAGGCTGAGCAAGAAATGATTTCTGAGAGATCA 4500
Qy 4501 AAAGTCTCATTTATGCAATGAAAGCCAGAGAGCTGGAACAGAGCCGATGAAATTT 4560
Db 4501 AAAGTCTCATTTATGCAATGAAAGCCAGAGAGCTGGAACAGAGCCGATGAAATTT 4560
Qy 4561 GAGCTGTGCTTCCGAGCGGAGATGATCTATTGATGATGCTGCTGCTTCCGAACTC 4620
Db 4561 GAGCTGTGCTTCCGAGCGGAGATGATCTATTGATGATGCTGCTGCTTCCGAACTC 4620
Qy 4621 GCAAAATCAAGCAAGCAAGATGTCCATCACTGAAAGATGAAATCTCACCCGACACC 4680
Db 4621 GCAAAATCAAGCAAGCAAGATGTCCATCACTGAAAGATGAAATCTCACCCGACACC 4680
Qy 4681 ACCGCTGAGCCCGGAGAGACCTCTACTGTAGCTCCAGCTTCCGTAACAACAGGCG 4740
Db 4681 ACCGCTGAGCCCGGAGAGACCTCTACTGTAGCTCCAGCTTCCGTAACAACAGGCG 4740
Qy 4741 TGGGTCAACGCTTGAATCAAGTGTGCGCAGGTGGAGAGTTTCTAGGAAAAAGCAGAA 4800

Db 4741 TGGGTCAACGCTTGAATCAAGTGTGCGCAGGTGGAGAGTTTCTAGGAAAAAGCAGAA 4800
Qy 4801 GCTGATGCTAAATCTGCTTGGAAAATCTCCCTGTGAACCTGGAAGTGAAGCCTGTAGAC 4860
Db 4801 GCTGATGCTAAATCTGCTTGGAAAATCTCCCTGTGAACCTGGAAGTGAAGCCTGTAGAC 4860
Qy 4861 ATGAACTGCAAGCTGCTTCACTGACCAAGTGTGTGTTGGGCAACCGAGAAAGGGCTC 4920
Db 4861 ATGAACTGCAAGCTGCTTCACTGACCAAGTGTGTGTTGGGCAACCGAGAAAGGGCTC 4920
Qy 4921 TACGCCCTGAATGTCTTGAAGAAATCTCCCTAACCAATGTCAGAGAAATGGAGCTTTC 4980
Db 4921 TACGCCCTGAATGTCTTGAAGAAATCTCCCTAACCAATGTCAGAGAAATGGAGCTTTC 4980
Qy 4981 CAATTTATTTATCAAGGACTGAGAAAGCTA CTCATGATAGCAGAGAAAGCGGCA 5040
Db 4981 CAATTTATTTATCAAGGACTGAGAAAGCTA CTCATGATAGCAGAGAAAGCGGCA 5040
Qy 5041 CTGTGCTTGTGAGCTGAAAGAAAGTGAACATGTCCTGTGGCCAGTCCCACTGCGCTGCC 5100
Db 5041 CTGTGCTTGTGAGCTGAAAGAAAGTGAACATGTCCTGTGGCCAGTCCCACTGCGCTGCC 5100
Qy 5101 CAGCCGACATCTCACCCCAATTTTGAAGCTGTCAAGGGCTGCCACTTGTGAGGCA 5160
Db 5101 CAGCCGACATCTCACCCCAATTTTGAAGCTGTCAAGGGCTGCCACTTGTGAGGCA 5160
Qy 5161 GGCAGATTTGAGAAAGGGCTCTGATCTGTGAGCAAGCCATGCCAGCAAAAGTGTCTC 5220
Db 5161 GGCAGATTTGAGAAAGGGCTCTGATCTGTGAGCAAGCCATGCCAGCAAAAGTGTCTC 5220
Qy 5221 CGCTACAGAAATCTCAGCAAAATCTGCAATCCGGAAGATGAGACTCAGAGCC 5280
Db 5221 CGCTACAGAAATCTCAGCAAAATCTGCAATCCGGAAGATGAGACTCAGAGCC 5280
Qy 5281 TGGAGCTGTATCCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCA 5340
Db 5281 TGGAGCTGTATCCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCA 5340
Qy 5341 ATGACATGAAAGAGTACACGCTGAGAGAAATCTGATTAAGATGCAATTCCTTGGCA 5400
Db 5341 ATGACATGAAAGAGTACACGCTGAGAGAAATCTGATTAAGATGCAATTCCTTGGCA 5400
Qy 5401 CTTGCTGTGTTGCGGCTCTTCCAAAGCTTCCCTGTCTCAATCTGTCAGGTGAACAGC 5460
Db 5401 CTTGCTGTGTTGCGGCTCTTCCAAAGCTTCCCTGTCTCAATCTGTCAGGTGAACAGC 5460
Qy 5461 GCAAGGAGAGAGAGAGTACTTGTGTGTTTCCAGAAATTTGAGAGTGTGAGATTTCT 5520
Db 5461 GCAAGGAGAGAGAGAGTACTTGTGTGTTTCCAGAAATTTGAGAGTGTGAGATTTCT 5520
Qy 5521 TACGGAAGAGTGAAGCCGACAGAGATCTCAAGTGAAGTGTGCTTACCTTTGGCCTTGGC 5580
Db 5521 TACGGAAGAGTGAAGCCGACAGAGATCTCAAGTGAAGTGTGCTTACCTTTGGCCTTGGC 5580
Qy 5581 TACGGAAGAGTGAAGTGTGTTGTAAGCCACTTCACTCACTGAAATGAGATCAAG 5640
Db 5581 TACGGAAGAGTGAAGTGTGTTGTAAGCCACTTCACTCACTGAAATGAGATCAAG 5640
Qy 5641 GCAAGCTCTCAAG 5700
Db 5641 GCAAGCTCTCAAG 5700
Qy 5701 CTGGGCTTGTGCAATTTCTGAGAGAGATTTGAGAGTGTGCTGATACAGAGATTAATTA 5760
Db 5701 CTGGGCTTGTGCAATTTCTGAGAGAGATTTGAGAGTGTGCTGATACAGAGATTAATTA 5760
Qy 5761 AAGGTCAATTTGCTGCAAG 5820
Db 5761 AAGGTCAATTTGCTGCAAG 5820
Qy 5821 CCGTCAACCTTCCGAG 5887
Db 5821 CCGTCAACCTTCCGAG 5887

RESULT 13
AX429514
LOCUS AX429514 6159 bp DNA linear PAT 21-JUN-2002
DEFINITION Sequence 3 from Patent WO0234896.
ACCESSION AX429514
VERSION AX429514.1 GI:21540792
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Bukaryaota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS Kapeller-Libermann, R.
TITLE 13245, a novel human myotonic dystrophy type protein kinase and
uses thereof
JOURNAL Patent: WO 0234896-A 3 02-MAY-2002;
MILLENNIUM PHARM INC (US)
FEATURES
source location/Qualifiers
1. 6159
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 91.8%; Score 5661.4; DB 6; Length 6159;
Best Local Similarity 95.8%; Pred. No. 0;
Matches 5944; Conservative 0; Mismatches 16; Indels 243; Gaps 3;
1 ANGTTGAAGTTCAAAATATGAGGCGGGAATCCTTGGATGCTGCTGTGAACCATTT
1 AAGTTGAAGTTCAAAATATGAGGCGGGAATCCTTGGATGCTGCTGTGAACCATTT
60
61 GCCAGCCGGGCTTCAGAGCTGAATCTGTTCTTCAGGGGAAACACCTTTATGACTCAA
61 GCCAGCCGGGCTTCAGAGCTGAATCTGTTCTTCAGGGGAAACACCTTTATGACTCAA
120
121 CAGCAGATGCTCTCTCTTCCGGAAGAGGATTTAGATGCCCTCTTGTCTCTTTGAA
121 CAGCAGATGCTCTCTCTTCCGGAAGAGGATTTAGATGCCCTCTTGTCTCTTTGAA
180
121 GAATGAGTCAGGCTGCTGATGAAGATTAAAGCATGAGCACTTTGTCGGAAGTAT
181 GAATGAGTCAGGCTGCTGATGAAGATTAAAGCATGAGCACTTTGTCGGAAGTAT
240
181 GAATGAGTCAGGCTGCTGATGAAGATTAAAGCATGAGCACTTTGTCGGAAGTAT
240
241 TCCGACACCATAGCTGAGTTACAGAGCTTCAGCCTTCGCAAAAGGACTTCAGATCAGA
241 TCCGACACCATAGCTGAGTTACAGAGCTTCAGCCTTCGCAAAAGGACTTCAGATCAGA
300
301 AGTCTTGATGTTGTGCTCACTTGTCTGAAGTGCAGTGGTAAAGAGAAAGCAACCGGG
301 AGTCTTGATGTTGTGCTCACTTGTCTGAAGTGCAGTGGTAAAGAGAAAGCAACCGGG
360
301 AGTCTTGATGTTGTGCTCACTTGTCTGAAGTGCAGTGGTAAAGAGAAAGCAACCGGG
360
361 GACATCTATGCTATGAAGATGAAGAGGCTTTATGGCCCAAGAGAGGTTTCA
361 GACATCTATGCTATGAAGATGAAGAGGCTTTATGGCCCAAGAGAGGTTTCA
420
421 TTTTGTGAGAGAGAGGAGCAATATTAATCTCGAAGCAAGCCCGGATGCCCAATTA
421 TTTTGTGAGAGAGAGGAGCAATATTAATCTCGAAGCAAGCCCGGATGCCCAATTA
480
481 CAGTATGCTTTTCAGAGCAAAAATCACTTTATCTGATGAGGAATATCAGCTGAGGG
481 CAGTATGCTTTTCAGAGCAAAAATCACTTTATCTGATGAGGAATATCAGCTGAGGG
540
481 CAGTATGCTTTTCAGAGCAAAAATCACTTTATCTGATGAGGAATATCAGCTGAGGG
540
541 GACTTGCTGCTCACTTTTGAATAGATATGAGGACGAGTTAGATGAAGAACTGATCACTTT
541 GACTTGCTGCTCACTTTTGAATAGATATGAGGACGAGTTAGATGAAGAACTGATCACTTT
600
541 GACTTGCTGCTCACTTTTGAATAGATATGAGGACGAGTTAGATGAAGAACTGATCACTTT
600
601 TACCTAGCTGAGCTGATTTGGCTGTTCAACAGCTTCATCTGATGGGATAGCTGATCAGA
601 TACCTAGCTGAGCTGATTTGGCTGTTCAACAGCTTCATCTGATGGGATAGCTGATCAGA
660
601 TACCTAGCTGAGCTGATTTGGCTGTTCAACAGCTTCATCTGATGGGATAGCTGATCAGA
660

QY 661 GACATCAAGCCTGAGAAATTTCTGTTGACCGGACAGAGACATCAAGCTGTGGATTTT 720
DB 661 GACATCAAGCCTGAGAAATTTCTGTTGACCGGACAGAGACATCAAGCTGTGGATTTT 720
QY 721 GGAATCGCGCGGAAATGATTCAGAACAGATGCTGAAATGCGCAACCTCCGATTTGGGACC 780
DB 721 GGAATCGCGCGGAAATGATTCAGAACAGATGCTGAAATGCGCAACCTCCGATTTGGGACC 780
QY 781 CCAGATTACATGCTCCTGTAAGTCTGATCTGTATGAAACGGGATGAGAAAGGACCTTAC 840
DB 781 CCAGATTACATGCTCCTGTAAGTCTGATCTGTATGAAACGGGATGAGAAAGGACCTTAC 840
QY 841 GGCCTGAGCTGTACTGTGTGCTGAGTGGGCGTATGCTTATGATGATTTATGAGAGA 900
DB 841 GGCCTGAGCTGTACTGTGTGCTGAGTGGGCGTATGCTTATGATGATTTATGAGAGA 900
QY 901 TCCGCTGAGCTGAGTGTGCTGAGTGGGCGTATGCTTATGATGATTTATGAGAGA 960
DB 901 TCCGCTGAGCTGAGTGTGCTGAGTGGGCGTATGCTTATGATGATTTATGAGAGA 960
QY 961 TTTTGAATAATTCAGATGACCCCAAGTGAAGCTGATCTTCTTATGATCTTCAAGC 1020
DB 961 TTTTGAATAATTCAGATGACCCCAAGTGAAGCTGATCTTCTTATGATCTTCAAGC 1020
QY 1021 TTGTTGTGCGGCGGAGAGAGAGAGAGAGAGTGAAGGTTTGTGCTGCAATCCTTTCTTC 1080
DB 1021 TTGTTGTGCGGCGGAGAGAGAGAGAGAGAGTGAAGGTTTGTGCTGCAATCCTTTCTTC 1080
QY 1081 TCTTAAATGATGCTGAGCAACATTCGTAACTCTCTCTCCCTCCCTGCTTCCACCTTCAG 1140
DB 1081 TCTTAAATGATGCTGAGCAACATTCGTAACTCTCTCTCCCTCCCTGCTTCCACCTTCAG 1140
QY 1141 TCTGAGATGACACCTCCATTTTGTATGAACAGAGAAATTCGTGGGTTTCACTCTCT 1200
DB 1141 TCTGAGATGACACCTCCATTTTGTATGAACAGAGAAATTCGTGGGTTTCACTCTCT 1200
QY 1201 CCGTGCAGCTGAGCCCTTCAGGCTTCTGCGGTGAAGAACTGCGGTTTGTGGGTTTTCG 1260
DB 1201 CCGTGCAGCTGAGCCCTTCAGGCTTCTGCGGTGAAGAACTGCGGTTTGTGGGTTTTCG 1260
QY 1261 TACAGAGAGGCACTGGGGATTTCTGTGATATCTGATCTGTGTGTGCTGAGCTTC 1320
DB 1261 TACAGAGAGGCACTGGGGATTTCTGTGATATCTGATCTGTGTGTGCTGAGCTTC 1320
QY 1321 CCTGCAAGACTAGCTCCATGAGAAAGAACTTCTCATCAAAAGCAAGGCTTCAAGAC 1380
DB 1321 CCTGCAAGACTAGCTCCATGAGAAAGAACTTCTCATCAAAAGCAAGGCTTCAAGAC 1380
QY 1381 TCTCAAGACAACTGTCACAAGATGAGAGAGAAATGACCCGGTTTATCTGAGAGTGTCA 1440
DB 1381 TCTCAAGACAACTGTCACAAGATGAGAGAGAAATGACCCGGTTTATCTGAGAGTGTCA 1440
QY 1441 GAGGTGAGAGGCTGTCTTATGATGAAGAGGAGGAGGCTGAGAGCTTCAAGAGCTCAAGAGA 1500
DB 1441 GAGGTGAGAGGCTGTCTTATGATGAAGAGGAGGAGGCTTCAAGAGCTTCAAGAGA 1500
QY 1501 TCCCTCTGAGAGCAGAGCCTTGTCTTCAATCAAGATGAGTGAAGTTAAAGGAGT 1560
DB 1501 TCCCTCTGAGAGCAGAGCCTTGTCTTCAATCAAGATGAGTGAAGTTAAAGGAGT 1560
QY 1561 TTGAGAGAGCAGAGAGGAGTGTCCAGAGAGATGACAAAGCACTGAGCTTCTCAT 1620
DB 1561 TTGAGAGAGCAGAGAGGAGTGTCCAGAGAGATGACAAAGCACTGAGCTTCTCAT 1620
QY 1621 GATATCAAGAGCAGAGAGGAGTGTCCAGAGAGATGACAAAGCACTGAGCTTCA 1680
DB 1621 GATATCAAGAGCAGAGAGGAGTGTCCAGAGAGATGACAAAGCACTGAGCTTCA 1680
QY 1681 GTGAGAGAAATGAGAGTGTATGATGAATCAGTTGAAAGAGATCTTGTCTCAGCAAGAGAA 1740
DB 1681 GTGAGAGAAATGAGAGTGTATGATGAATCAGTTGAAAGAGATCTTGTCTCAGCAAGAGAA 1740

QY 1741 CGGAGTGTCTTACGAAATTGAGCTGAGAGAGTCTGCGCTTGCTGTGAAGAATTCAAG 1800
DB 1741 CGGAGGATCTTACGAATCTGAGAGAGTCTGCGCTTGCTGTGAAGAATTCAAG 1800
QY 1801 CGGAAAGCGACAGAAATGTCAGCAATACTGTTGAAGCTAAAGATCAAGGAAAGCCCTGAA 1860
DB 1801 CGGAAAGCGACAGAAATGTCAGCAATACTGTTGAAGCTAAAGATCAAGGAAAGCCCTGAA 1860
QY 1861 GTGGGGAATATGCGAACTGGAAGAGTCAATGCTGAGCAGCAGCTCAAAATTCAAGAG 1920
DB 1861 GTGGGGAATATGCGAACTGGAAGAGTCAATGCTGAGCAGCAGCTCAAAATTCAAGAG 1920
QY 1921 CTCGAAAGAGAACTGGAAGAGCTGTAAAGCCAGCAGAGGCCACCGAGCTGCTGAG 1980
DB 1921 CTCGAAAGAGAACTGGAAGAGCTGTAAAGCCAGCAGAGGCCACCGAGCTGCTGAG 1980
QY 1981 AATATCCGACGCGAAAGAGCGGAGGAGGAGCTGGAAGAGCTGCAAGACCGAGAG 2040
DB 1945 -----GCGAAAGAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1992
QY 2041 GATTCTTCTGAAAGCATCAGAAAGAGCTGTGGAAGCTGAGGAAAGCCGCCATTCTCTG 2100
DB 1993 GATTCTTCTGAAAGCATCAGAAAGAGCTGTGGAAGCTGAGGAAAGCCGCCATTCTCTG 2052
QY 2101 GAGAACAGGTAAAGAGACTAGAGACCATGAGCGTGAAGAAAACAGACTGAAGATGAC 2160
DB 2053 GAGAACAGGTAAAGAGACTAGAGACCATGAGCGTGAAGAAAACAGACTGAAGATGAC 2112
QY 2161 ATCCAGACAAATCCCAACGATCCAGCAATGCTGATTAATTTCTGAGCTCGAAGAG 2220
DB 2113 ATCCAGACAAATCCCAACGATCCAGCAATGCTGATTAATTTCTGAGCTCGAAGAG 2172
QY 2221 AAACATGCGGAGGCCCAAGTCTCAGCCAGCACTAGAAGTGCACCTGAAGACAGAAAGAG 2280
DB 2173 AAACATGCGGAGGCCCAAGTCTCAGCCAGCACTAGAAGTGCACCTGAAGACAGAAAGAG 2232
QY 2281 CAGCACTATGAGGAAAAGATTAAAGTTGTTGAGCAATCAGATTAAGAAAAGCTGGCTGAC 2340
DB 2233 CAGCACTATGAGGAAAAGATTAAAGTTGTTGAGCAATCAGATTAAGAAAAGCTGGCTGAC 2292
QY 2341 AAGGAGACACTGAGGAACTGATGACAGACAGAGAGAGGCCCTGAGAAAGGGTAA 2400
DB 2293 AAGGAGACACTGAGGAACTGATGACAGACAGAGAGAGGCCCTGAGAAAGGGTAA 2352
QY 2401 ATTCTCAGGCAAGAGAGGATGATCAATGCTATGGAATTCAGAAATCAGATCCCTGAA 2460
DB 2353 ATTCTCAGGCAAGAGAGGATGATCAATGCTATGGAATTCAGAAATCAGATCCCTGAA 2412
QY 2461 CAGAGATTTGTGAACTGTCTGAAGCCAAATTAACCTTGACAAATAGCAGTCTTTTACC 2520
DB 2413 CAGAGATTTGTGAACTGTCTGAAGCCAAATTAACCTTGACAAATAGCAGTCTTTTACC 2472
QY 2521 CAAAGGAACTGAAGGCCCAAGAAAGATGATTTCTGAATCTCAGGCAACGAAATTTTAC 2580
DB 2473 CAAAGGAACTGAAGGCCCAAGAAAGATGATTTCTGAATCTCAGGCAACGAAATTTTAC 2532
QY 2581 CTGGAGACACAGGCTGGAGAGTTGGAAGGCCAGAACCGAAACTGAGGAGGAGCTGGAG 2640
DB 2533 CTGGAGACACAGGCTGGAGAGTTGGAAGGCCAGAACCGAAACTGAGGAGGAGCTGGAG 2592
QY 2641 AAGATCAGCCACCAAGACCAAGTGAACAGATTCGAGTCTGGAACCTGAGAGCAAGATTG 2700
DB 2593 AAGATCAGCCACCAAGACCAAGTGAACAGATTCGAGTCTGGAACCTGAGAGCAAGATTG 2652
QY 2701 CGGGAAGTCACTGAGCAAGAGAGCAAGAACTGGAAGTCAAGGCCGCTCAAGAG 2760
DB 2653 CGGGAAGTCACTGAGCAAGAGAGCAAGAACTGGAAGTCAAGGCCGCTCAAGAG 2712
QY 2761 CTACAGCTCTCCGCGAGGAGCGGAGTCAAGTTGACAGCCCTGAGGCTGACCGGCGG 2820
DB 2713 CTACAGCTCTCCGCGAGGAGCGGAGTCAAGTTGACAGCCCTGAGGCTGACCGGCGG 2772
QY 2821 GCCCTGAGAGCAAGCTTGCAGAGCGAAGACAGAGCTGGAAGAGCAACAGCAGAAAGCT 2880

DB 2773 GCCCTGAGAGCAAGCTTGCAGAGCGAAGACAGAGCTGGAAGAGCAACAGCAAGAGCT 2832
QY 2881 GAAGAGAGATCCAGGCACTCAGGACATAGAGATGAATAATCCAGGCCAAATTTGATGCT 2940
DB 2833 GAAGAGAGATCCAGGCACTCAGGCACTAGAGATGAATAATCCAGGCCAAATTTGATGCT 2892
QY 2941 CTTGTAAACAGCTGTATCTGTATATCAAGACTGGAAGAGCAGCTTAAACAGCTGACCGAG 3000
DB 2893 CTTGTAAACAGCTGTATCTGTATATCAAGACTGGAAGAGCAGCTTAAACAGCTGACCGAG 2952
QY 3001 GACAAAGCTGAATCAACCAACCAAACTTCTACTTGTCCAAACAATCCGATGAGGCTTCT 3060
DB 2953 GACAAAGCTGAATCAACCAACCAAACTTCTACTTGTCCAAACAATCCGATGAGGCTTCT 3012
QY 3061 GCGGCCACGACGAGATTGTACAACTGCGAAAGTGAAGTGAACCATCTCCGCGGAGATC 3120
DB 3013 GCGGCCACGACGAGATTGTACAACTGCGAAAGTGAAGTGAACCATCTCCGCGGAGATC 3072
QY 3121 ACGGAAAGAGATGACAGCTTACCAAGCCAGAACCAACGATGAGGCTCTGAAGACCAAG 3180
DB 3073 ACGGAAAGAGATGACAGCTTACCAAGCCAGAACCAACGATGAGGCTCTGAAGACCAAG 3132
QY 3181 TGCACATGCTGAGAGGAAACAGGTATGATTTGGAAGCCCTTAAACGATGAGCTGTGAA 3240
DB 3133 TGCACATGCTGAGAGGAAACAGGTATGATTTGGAAGCCCTTAAACGATGAGCTGTGAA 3192
QY 3241 AAGAGCGGACAGGAGGAGGCTGAGAGAGGCTCTGAGGATGATGAAGAAATCCAGTTTGA 3300
DB 3193 AAGAGCGGACAGGAGGAGGCTGAGAGAGGCTCTGAGGATGATGAAGAAATCCAGTTTGA 3252
QY 3301 TGTGCGGTTTGAAGGCTGAGAGGATGCTGGAACCGAGAAACAGACAGAGGCGAGAGCC 3360
DB 3253 TGTGCGGTTTGAAGGCTGAGAGGATGCTGGAACCGAGAAACAGACAGAGGCGAGAGCC 3312
QY 3361 GATCAGCGATCAACCGATCTCCGCAAGTGTGAGCTGACATGAAAGACACAGGCT 3420
DB 3313 GATCAGCGATCAACCGATCTCCGCAAGTGTGAGCTGACATGAAAGACACAGGCT 3372
QY 3421 GAGATTTCTGCTGAGAGGAGGCTCTCAAGAGACAGAGCTGAAGGCCGAGAGCTCTCT 3480
DB 3373 GAGATTTCTGCTGAGAGGAGGCTCTCAAGAGACAGAGCTGAAGGCCGAGAGCTCTCT 3432
QY 3481 GACAACTCAATGACCTGAGAAAGAGCAATGCTATGTTGAATGAATGCCCGAAGCTTA 3540
DB 3433 GACAACTCAATGACCTGAGAAAGAGCAATGCTATGTTGAATGAATGCCCGAAGCTTA 3492
QY 3541 CAGCAGAGCTGAGACTGAAGAGAGCTCAACAGAGCTTCTGAAAGCAAGCCAA 3600
DB 3493 CAGCAGAGCTGAGACTGAAGAGAGCTCAACAGAGCTTCTGAAAGCAAGCCAA 3552
QY 3601 TTACAGCAGAGTGAACCTGCAAGAAATCAATTTCCGCTGACTCAAGACTGCA 3660
DB 3553 TTACAGCAGAGTGAACCTGCAAGAAATCAATTTCCGCTGACTCAAGACTGCA 3612
QY 3661 GAAGCTCTAGATGGGCTGATCTAGTGAAGACAGAAAGAGTACTTGAAGTACAGCTG 3720
DB 3613 GAAGCTCTAGATGGGCTGATCTAGTGAAGACAGAAAGAGTACTTGAAGTACAGCTG 3672
QY 3721 GAAACATTCAGGTTCTATTTCTATGAAAGGTGAATAATGGAAGCACTATTTCTCA 3780
DB 3673 GAAACATTCAGGTTCTATTTCTATGAAAGGTGAATAATGGAAGCACTATTTCTCA 3732
QY 3781 CAAACCAACTCTATGATTTTCTGCAAGCCAAATGGAACCACTGTGTAAGAAAGAA 3838
DB 3733 CAAACCAACTCTATGATTTTCTGCAAGCCAAATGGAACCACTGTGTAAGAAAGAA 3792
QY 3839 -----AGTTCTCTGCACTAC 3855
DB 3793 GGTTTATTTAGTGAAGAAAGAGACCTGCTTACCAACAGAGTTCTCTGCACTAC 3852
QY 3856 AATGAGCTGAAGCTGAGCCTGAGAGAAAGAGAAAGCTGCTGTCAGAGCTTGAAGAAAGCC 3915

Dh 3853 AATGAGCTGAAGCTGGCCCTGGAGAGAGAAAGCTCGCTGTGCAAGCTTAGAGAAAGCC 3912
Qy 3916 CTTTCAGAAAGACCCGCAATGAGAGCTCCGCTCCGCCGGAGAGAAAGTGTCCACCCGAAAGCA 3975
Dh 3913 CTTTCAGAAAGACCCGCAATGAGAGCTCCGCTCCGCCGGAGAGAAAGTGTCCACCCGAAAGCA 3972
Qy 3976 ACCGAGCAACCAACCAATCCAGCCGAGCCACCGCGAGAGAGAGATTCGCAATGTCGGCC 4035
Dh 3973 ACCGAGCAACCAACCAATCCAGCCGAGCCACCGCGAGAGAGAGATTCGCAATGTCGGCC 4032
Qy 4036 ATCGTGGGTGCGCCAGAGACACAGCCCAATGAGCTGTGAGCCCGCCATCCAGC 4095
Dh 4033 ATCGTGGGTGCGCCAGAGACACAGCCCAATGAGCTGTGAGCCCGCCATCCAGC 4092
Qy 4096 CGAGAAAGAGATCTTCAATCTCCAGAGAAATTTAGTGGCGTCTTAAAGAACGCATGCAC 4155
Dh 4093 CGAGAAAGAGATCTTCAATCTCCAGAGAAATTTAGTGGCGTCTTAAAGAACGCATGCAC 4152
Qy 4156 CACAAATATTTCTGACCGAATTCAGGTAGGACTGAGGACCAAAAGTGTGCTGTG 4215
Dh 4153 CACAAATATTTCTGACCGAATTCAGGTAGGACTGAGGACCAAAAGTGTGCTGTG 4212
Qy 4216 TGTCTGGAATACCGTGAACCTTGGAGCGCAGGCAATCCAAATGTCTCGAATGTCAAGTATG 4275
Dh 4213 TGTCTGGAATACCGTGAACCTTGGAGCGCAGGCAATCCAAATGTCTCGAATGTCAAGTATG 4272
Qy 4276 TGTCAACCCCAAGTGTCTCACGTGCTTGCACGCCACTGCGGCTTGCTGCTGAATATGCC 4335
Dh 4273 TGTCAACCCCAAGTGTCTCACGTGCTTGCACGCCACTGCGGCTTGCTGCTGAATATGCC 4332
Qy 4336 ACACACTTTCACCGAGGCTTCTGCGGTGACAAATGAACTCCCAAGTGTCTCCAGACCAAG 4395
Dh 4333 ACACACTTTCACCGAGGCTTCTGCGGTGACAAATGAACTCCCAAGTGTCTCCAGACCAAG 4392
Qy 4396 GAGCCGAGAGAGAGCTTGGACCTTGAAGAGGTGATGAAGTGTGCCAGAAATACAAACGA 4455
Dh 4393 GAGCCGAGAGAGAGCTTGGACCTTGAAGAGGTGATGAAGTGTGCCAGAAATACAAACGA 4452
Qy 4456 GGAACAGAGAGGTGGAGACAGAAATGATGTCCTGAGAGGATCAAAAGTCTCATTTAT 4515
Dh 4453 GGAACAGAGAGGTGGAGACAGAAATGATGTCCTGAGAGGATCAAAAGTCTCATTTAT 4512
Qy 4516 GACAAATGAAGCCAGAGAGCTGGACAGAGCCGCTGAGAAATTTAGAGCTGTGCTTCCC 4575
Dh 4513 GACAAATGAAGCCAGAGAGCTGGACAGAGCCGCTGAGAAATTTAGAGCTGTGCTTCCC 4572
Qy 4576 GAGCGGAGATGATCTATTCATGAGTGGCTGTGAGTCTTCCGAATCTGCAAAATACAGCCAA 4635
Dh 4573 GAGCGGAGATGATCTATTCATGAGTGGCTGTGAGTCTTCCGAATCTGCAAAATACAGCCAA 4632
Qy 4636 GCAAGATGTCCTCAATCACTAATGAGATGAAATCTCACCCGCAACCACTGTGGCCCGGG 4695
Dh 4633 GCA----- 4635
Qy 4696 AGAACCTCTACTGTGCTAGCTCCAGCTTCCCTGACAAACAGCGCTGGGTCAACCGCTTA 4755
Dh 4636 ----- 4635
Qy 4756 GAATCAGTTGTGCAAGTGGAGAGTTTCTAGGAGAAAGAGAGCTGATGTAACCTG 4815
Dh 4636 -----GAAAAAGCAGAACTGATGTAACCTG 4662
Qy 4816 CTTGGAATCTCCCTGCTGAAACTGAGAGGTGATGACCGTCTAGACATGAACTGCAGCTG 4875
Dh 4663 CTTGGAATCTCCCTGCTGAAACTGAGAGGTGATGACCGTCTAGACATGAACTGCAGCTG 4722
Qy 4876 CCGTTCAAGTGAACAGTGTGTGTTGGTGGGACCCGAGAAAGGCTCTAGCGCCCTGAATGTC 4935
Dh 4733 CCGTTCAAGTGAACAGTGTGTGTTGGTGGGACCCGAGAAAGGCTCTAGCGCCCTGAATGTC 4782
Qy 4936 TTGAAAAACTCCCTAACCATGTCCCAAGAAATTGGAGAGTCTTCAAAATTTATTTATC 4995
Dh 4783 TTGAAAAACTCCCTAACCATGTCCCAAGAAATTGGAGAGTCTTCAAAATTTATTTATC 4842

Qy 4996 AAGAACCTGAGAAAGCTACTCATGATAGCAGAGAAAGAGCGGCACTGTGTCTTGTGAC 5055
Dh 4843 AAGAACCTGAGAAAGCTACTCATGATAGCAGAGAAAGAGCGGCACTGTGTCTTGTGAC 4902
Qy 5056 GTGAAGAAAGTGAACAGTGTCCCTGGCCAGTCCCACTGTCTTCCAGCCGACATTTCA 5115
Dh 4903 GTGAAGAAAGTGAACAGTGTCCCTGGCCAGTCCCACTGTCTTCCAGCCGACATTTCA 4962
Qy 5116 CCCAATATTTTGAAGCTGTCAAGGGCTGCACTTGTGTTGGGGCAGGCAAGATTTGAGAAC 5175
Dh 4963 CCCAATATTTTGAAGCTGTCAAGGGCTGCACTTGTGTTGGGGCAGGCAAGATTTGAGAAC 5022
Qy 5176 GGGCTCTGCAATCTGTGCAAGCCATGCCCAGAAAGTGTGATTTCTCCGCTACAAAGAAAC 5235
Dh 5023 GGGCTCTGCAATCTGTGCAAGCCATGCCCAGAAAGTGTGATTTCTCCGCTACAAAGAAAC 5082
Qy 5236 CTCAGCAATATCTGCAATCCGGAAGAGATGAGAACTCGAAGCCCTGAGCTGTATCCAC 5295
Dh 5083 CTCAGCAATATCTGCAATCCGGAAGAGATGAGAACTCGAAGCCCTGAGCTGTATCCAC 5142
Qy 5296 TTCACCAATTTACAGTATCTCTCATTTGGAAACCAATAAATTTCTACGAATTCACATGAAGAC 5355
Dh 5143 TTCACCAATTTACAGTATCTCTCATTTGGAAACCAATAAATTTCTACGAATTCACATGAAGAC 5202
Qy 5356 TACAGCTGAGAGAAATTCCTGGAATGAAGATGACCAATTCCTTGGCACTGTGTTTGGC 5415
Dh 5203 TACAGCTGAGAGAAATTCCTGGAATGAAGATGACCAATTCCTTGGCACTGTGTTTGGC 5262
Qy 5416 GCTCTTCCAAACAGTCTCCCTGTCTCAATCTGTGAGAGTGAACAGCCGAGGACGAGAG 5475
Dh 5263 GCTCTTCCAAACAGTCTCCCTGTCTCAATCTGTGAGAGTGAACAGCCGAGGACGAGAG 5322
Qy 5476 GAGTACTTGTCTGTGTTCCAGCAATTTGGAGGTGTGATGTTCTTACGGAAGACGTAC 5535
Dh 5323 GAGTACTTGTCTGTGTTCCAGCAATTTGGAGGTGTGATGTTCTTACGGAAGACGTAC 5382
Qy 5536 GCGACAGAGATCTGAAGTGAAGTGTGCTTACCTTTGGCTTTCACAGAAACCTTAT 5595
Dh 5383 GCGACAGAGATCTGAAGTGAAGTGTGCTTACCTTTGGCTTTCACAGAAACCTTAT 5442
Qy 5596 CTGTTTGGACCCCACTTCAACTCATCTCGAAGTAAATGAGATTCAGAGCAGCTCTCAGCA 5655
Dh 5443 CTGTTTGGACCCCACTTCAACTCATCTCGAAGTAAATGAGATTCAGAGCAGCTCTCAGCA 5502
Qy 5656 GGGACCCCTGCGCGAGCGCTACCTGAGCAATCCGAAACCCGCGCTACCTGGAGCCCTGCAAT 5715
Dh 5503 GGGACCCCTGCGCGAGCGCTACCTGAGCAATCCGAAACCCGCGCTACCTGGAGCCCTGCAAT 5562
Qy 5716 TCCTCAGAGAGATTTACTTGGCGTCTTCATACAGATTAATTAAGGATCAATTTGCTGC 5775
Dh 5563 TCCTCAGAGAGATTTACTTGGCGTCTTCATACAGATTAATTAAGGATCAATTTGCTGC 5622
Qy 5776 AAGGAAATCTGTGTAAGAGTCCGCGCACTGAACACACCGGGGCCCTTCCACTTCCCGC 5835
Dh 5623 AAGGAAATCTGTGTAAGAGTCCGCGCACTGAACACACCGGGGCCCTTCCACTTCCCGC 5682
Qy 5836 AGCAAGCCCAACAAAGCGAGCCCAACCAAGTACAGAGACATCAACAAAGCGCTGGCC 5895
Dh 5683 AGCAAGCCCAACAAAGCGAGCCCAACCAAGTACAGAGACATCAACAAAGCGCTGGCC 5742
Qy 5896 TTCAGCCCAAGCGCGCGCGGAAAGGCCCAAGCAACCCGCGAGAGCCAAAGCAACCCACCGC 5955
Dh 5743 TTCAGCCCAAGCGCGCGCGGAAAGGCCCAAGCAACCCGCGAGAGCCAAAGCAACCCACCGC 5802
Qy 5956 TACCGCGAGGGCGGAGACCGAGCTGTGCGAGGACAAATGTTCTGTGCGCCCTTGGAGGA 6015
Dh 5803 TACCGCGAGGGCGGAGACCGAGCTGTGCGAGGACAAATGTTCTGTGCGCCCTTGGAGGA 5862
Qy 6016 GAGAAATCCCGCGCGGATACAGCAAGCGAGAGAGCGGATCCCGCGAGAGCGTGTTC 6075
Dh 5863 GAGAAATCCCGCGCGGATACAGCAAGCGAGAGAGCGGATCCCGCGAGAGCGTGTTC 5922

RESULT 14
AX429512
LOCUS AX429512 6574 bp DNA linear PAT 21-JUN-2002
DEFINITION Sequence 1 from Patent WO0234896.
ACCESSION AX429512
VERSION AX429512.1 GI:21540791
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1
AUTHORS Kapeller-Liebermann, R.
TITLE 13245, a novel human myotonic dystrophy type protein kinase and
uses thereof
JOURNAL Patent: WO 0234896-A 1 02-MAY-2002;
MILLENNIUM PHARM INC (US)
FEATURES
Source 1. 6574
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 91.8%; Score 5661.4; DB 6; Length 6574;
Best Local Similarity 95.8%; Pred. No. 0;
Matches 5944; Conservative 0; Mismatches 16; Indels 243; Gaps 3;
QY 1 ATGTTGAAGTTCAATATGAGCGCGAATCCTTGTGATGCTGTGCTGAACCAT 60
Db 19 ATGTTGAAGTTCAATATGAGCGCGAATCCTTGTGATGCTGTGCTGAACCAT 78
QY 61 GCCAGCGGCGCTTCAGAGCTGAATCTGTCTTCCAGGGGAAACACCTTTATGACTCAA 120
Db 79 GCCAGCGGCGCTTCAGAGCTGAATCTGTCTTCCAGGGGAAACACCTTTATGACTCAA 138
QY 121 CAGCAAGTGTCTCTCTTCCCGAAGGATATTAATGCCCTTTGTTCTTTGAA 180
Db 139 CAGCAAGTGTCTCTCTTCCCGAAGGATATTAATGCCCTTTGTTCTTTGAA 198
QY 181 GAATGACGTAGCTGTCTGATGAAGATTAAAGACGTAGCACTTGTCCGGAAGTAT 240
Db 199 GAATGACGTAGCTGTCTGATGAAGATTAAAGACGTAGCACTTGTCCGGAAGTAT 258
QY 241 TCCGACACCATAGCTGAATTACAGAGCTCCAGCTTCGCGAAGACCTTGAAGTACA 300
Db 259 TCCGACACCATAGCTGAATTACAGAGCTCCAGCTTCGCGAAGACCTTGAAGTACA 318
QY 301 AGTCTTGAAGTTGTGCTCACTTGTCTGAAGTGAAGTGTAAAGAAAGCAACCGGG 360
Db 319 AGTCTTGAAGTTGTGCTCACTTGTCTGAAGTGAAGTGTAAAGAAAGCAACCGGG 378
QY 361 GACATCTATGCTATGAAGTGAAGAAAGGCTTTATTTGGCCCGAGAGCAGTTTCA 420
Db 379 GACATCTATGCTATGAAGTGAAGAAAGGCTTTATTTGGCCCGAGAGCAGTTTCA 438
QY 421 TTTTGAAGAAAGCGGAACATATTAATCTGAAGCAACAGCCCGTGAATCCCAATTA 480
Db 439 TTTTGAAGAAAGCGGAACATATTAATCTGAAGCAACAGCCCGTGAATCCCAATTA 498
QY 481 CAGTATGCTTTGAGCAAAATCACTTTATCTGGTCAATGAATATCAAGCTTGAGGG 540
Db 499 CAGTATGCTTTGAGCAAAATCACTTTATCTGGTCAATGAATATCAAGCTTGAGGG 558

QY 541 GACTGTGCTGACTTTTGAATAGATATGAGGACCACTTAATGAAAACCTGATACAGTTT 600
Db 559 GACTGTGCTGACTTTTGAATAGATATGAGGACCACTTAATGAAAACCTGATACAGTTT 618
QY 601 TACTAGTGAAGCTGATTTTGGCTTCAAGGGTTCATGATGGGATTCGTGATCGA 660
Db 619 TACTAGTGAAGCTGATTTTGGCTTCAAGGGTTCATGATGGGATTCGTGATCGA 678
QY 661 GACATCAAGCCTGAGAACTATTCGTGACCGGACAGAGACATCAAGCTGTGATTTT 720
Db 679 GACATCAAGCCTGAGAACTATTCGTGACCGGACAGAGACATCAAGCTGTGATTTT 738
QY 721 GGAATCTGCGGCAAAATGAATTAACAAGATGATGCAAACTCCGATTTGGAGC 780
Db 739 GGAATCTGCGGCAAAATGAATTAACAAGATGATGCAAACTCCGATTTGGAGC 798
QY 781 CCAATTACATGGCTCCCTGAAGTGTGACCTGTATGAACGGGGATGGAAGAAGCACTAC 840
Db 799 CCAATTACATGGCTCCCTGAAGTGTGACCTGTATGAACGGGGATGGAAGAAGCACTAC 858
QY 841 GGCCTGACTGTGACTGTGATGAGTGGGCGTGAATGCTATGAGATGATTTATGGAGA 900
Db 859 GGCCTGACTGTGACTGTGATGAGTGGGCGTGAATGCTATGAGATGATTTATGGAGA 918
QY 901 TCCCTCTGCGAAGGGAACCTTCGCAAACTTCAATTAATTAATTAATTTCAAGCG 960
Db 919 TCCCTCTGCGAAGGGAACCTTCGCAAACTTCAATTAATTAATTAATTTCAAGCG 978
QY 961 TTTTGAATTTTCCAGATGACCCCAAGTGAAGACGATCTTTGATCTGAATCAAGC 1020
Db 979 TTTTGAATTTTCCAGATGACCCCAAGTGAAGACGATCTTTGATCTGAATCAAGC 1038
QY 1021 TTGTTGTCGCGGCAAGAAAGAGACTGAAGTTGAAGTCTTTGCTCCATCTTTCTTC 1080
Db 1039 TTGTTGTCGCGGCAAGAAAGAGACTGAAGTTGAAGTCTTTGCTCCATCTTTCTTC 1098
QY 1081 TCTAAATTTGACTGGAACCAATTCGTACTCTCCCTTCCTTGTCCACCTCAAG 1140
Db 1099 TCTAAATTTGACTGGAACCAATTCGTACTCTCCCTTCCTTGTCCACCTCAAG 1158
QY 1141 TCTGACATGACACTTCCTCAATTTTGAATGAACAGAAAGAAATTCGTGGTTCATCTCT 1200
Db 1159 TCTGACATGACACTTCCTCAATTTTGAATGAACAGAAAGAAATTCGTGGTTCATCTCT 1218
QY 1201 CCGTGCAGCTGAGCCCTCAGGCTTCTCGGGTGAAGACCTGCGTTGTGGGGTTTTCG 1260
Db 1219 CCGTGCAGCTGAGCCCTCAGGCTTCTCGGGTGAAGACCTGCGTTGTGGGGTTTTCG 1278
QY 1261 TACAGCAAGGCACTGGGATTTCTGTGATCTGAGTCTGTGTGTCGGGTCTGAGCTCC 1320
Db 1279 TACAGCAAGGCACTGGGATTTCTGTGATCTGAGTCTGTGTGTCGGGTCTGAGCTCC 1338
QY 1321 CTGCGCAAGCTAGCTTCATGAAAAGAACTTCTCATCAAAAGCAAGACTTCAAGAC 1380
Db 1339 CTGCGCAAGCTAGCTTCATGAAAAGAACTTCTCATCAAAAGCAAGACTTCAAGAC 1398
QY 1381 TCTCAGGCAAGGTGCAAAAGTGAAGAGAGAGAAATGACCGGGTTCATCGGAGAGTCA 1440
Db 1399 TCTCAGGCAAGGTGCAAAAGTGAAGAGAGAGAAATGACCGGGTTCATCGGAGAGTCA 1458
QY 1441 GAGGTGAGGCTGTGCTTAAGTCAAGAGAGAGTGAAGGCTCTGAGACTCAGAGA 1500
Db 1459 GAGGTGAGGCTGTGCTTAAGTCAAGAGAGAGTGAAGGCTCTGAGACTCAGAGA 1518
QY 1501 TCCCTCTGAGAGAGACTTTGCTAATTAATCAAGAAATGCAATACCTTAAGCGGAAGT 1560
Db 1519 TCCCTCTGAGAGAGACTTTGCTAATTAATCAAGAAATGCAATACCTTAAGCGGAAGT 1578
QY 1561 TTGAGAGACAGAGAGAGAGAGTGTCCAGAGAGATGCAAAAGCACTGACCTTCTCAT 1620
Db 1579 TTGAGAGACAGAGAGAGAGTGTCCAGAGAGATGCAAAAGCACTGACCTTCTCAT 1638
QY 1621 GATATCAGAGAGCAGACCCGGAAGCTTCAAGAAATCAAGAGCAGAGTACAGAGCTCAA 1680

```
Db 1639 GATATCAGAGGCGAGCCGAGAGCTCCAGAAATCAAAGACGAGAGTACCAGGCTCA 1698
Qy 1681 GTGAAAGAAATGAGGTGATGATGAATCATGTTGGAAGAGATCTTTGCTCAGCAAGAA 1740
Db 1699 GTGAAAGAAATGAGGTGATGATGAATCATGTTGGAAGAGATCTTTGCTCAGCAAGAA 1758
Qy 1741 CGGAGTATCTCTAGATCTGAGCTGAGAGATCTCGGCTTGGCTGCTGAAGATTCAAG 1800
Db 1759 CGAGTATCTCTAGATCTGAGCTGAGAGATCTCGGCTTGGCTGCTGAAGATTCAAG 1818
Qy 1801 CGGAAAGCGAGAGATGTCAGCATATACTGTTGAAGCTAAGATCAAGGGAGCTGAA 1860
Db 1819 CGGAAAGCGAGAGATGTCAGCATATACTGTTGAAGCTAAGATCAAGGGAGCTGAA 1878
Qy 1861 GTGGAGAAATATGCGAAATCTGAGAAATCAATCTGAGCAGCTCAAAATTCAGAG 1920
Db 1879 GTGGAGAAATATGCGAAATCTGAGAAATCAATCTGAGCAGCTCAAAATTCAGAG 1938
Qy 1921 CTCGAAGAAATCTGAGAGGCTGTATAAGCCAGACCGAGGCCACCGAGCTGTCAG 1980
Db 1939 CTCGAAGAAATCTGAGAGGCT- 1962
Qy 1981 AATATCCGAGCGCAAGAGCGAGCGAGAGGAGCTGAGAAAGCTGCAAAACCGAG 2040
Db 1963 -----GCAAGAGCCAGCCGAGAGGAGCTGAGAAAGCTGCAAAACCGAG 2010
Qy 2041 GATTCTTCTGAAGGCATCAGAAAGAGCTGTGTGAGCTGAGAAACGCGCAATCTCTG 2100
Db 2011 GATTCTTCTGAAGGCATCAGAAAGAGCTGTGTGAGCTGAGAAACGCGCAATCTCTG 2070
Qy 2101 GAGAAAGGTTAAAGAGCTAGAGACCATAGAGCGTTAGAGAAACAGACTGAAGATGAC 2160
Db 2071 GAGAAAGGTTAAAGAGCTAGAGACCATAGAGCGTTAGAGAAACAGACTGAAGATGAC 2130
Qy 2161 ATCCAGACAAATATCCCAACAGATCCAGCAGATGAGCTGATAAATTTCTGAGCTCGAAG 2220
Db 2131 ATCCAGACAAATATCCCAACAGATCCAGCAGATGAGCTGATAAATTTCTGAGCTCGAAG 2190
Qy 2221 AAACATCGGAGGCCCAAGTCTCAGCCAGCACCCTAGAGTGCACTGAAACAGAAAGAG 2280
Db 2191 AAACATCGGAGGCCCAAGTCTCAGCCAGCACCCTAGAGTGCACTGAAACAGAAAGAG 2250
Qy 2281 CAGCATTAGAGGAAAGATTAAAGTGTGAGCAATCAAGTAAAGAAAGCCTGAGCTGAC 2340
Db 2251 CAGCATTAGAGGAAAGATTAAAGTGTGAGCAATCAAGTAAAGAAAGCCTGAGCTGAC 2310
Qy 2341 AAGGAGACACTGAGAGAACTGATGTCAGAGACAGAGAGAGAGGCCCATGAGAGAGGCAAA 2400
Db 2311 AAGGAGACACTGAGAGAACTGATGTCAGAGACAGAGAGAGAGGCCCATGAGAGAGGCAAA 2370
Qy 2401 ATTCTCAGCGAACAGAGGCGATGATCAATGCTATGAGATTCGAAGATCCAGATCCCTGAA 2460
Db 2371 ATTCTCAGCGAACAGAGGCGATGATCAATGCTATGAGATTCGAAGATCCAGATCCCTGAA 2430
Qy 2461 CAGAGATTTGGAACCTGCTGAGAGCAATTAACCTTGACCAATATGACGCTTTTACC 2520
Db 2431 CAGAGATTTGGAACCTGCTGAGAGCAATTAACCTTGACCAATATGACGCTTTTACC 2490
Qy 2521 CAAAGGACATGAGAGGCCCAAGAGAGATGATTCTGAACTCAGGCAACGAAATTTTAC 2580
Db 2491 CAAAGGACATGAGAGGCCCAAGAGAGATGATTCTGAACTCAGGCAACGAAATTTTAC 2550
Qy 2581 CTGAGAGACACAGGCTGGAAGTTGAGAGCCAGAACCGAAACTGAGAGAGCAGCTGAG 2640
Db 2551 CTGAGAGACACAGGCTGGAAGTTGAGAGCCAGAACCGAAACTGAGAGAGCAGCTGAG 2610
Qy 2641 AAGATCAAGCACACAGACCACTGTCAGAGATCGGCTGCTGGAATCTGAGACCAAGATTG 2700
Db 2611 AAGATCAAGCACACAGACCACTGTCAGAGATCGGCTGCTGGAATCTGAGACCAAGATTG 2670
Qy 2701 CGGAGAGTCACTGAGAGCAGAGGAGGAGAACTGAGAGCTCAAGGCCAGCTCAGAG 2760
Db 2671 CGGAGAGTCACTGAGAGCAGAGGAGGAGGAGAACTGAGAGCTCAAGGCCAGCTCAGAG 2730
Qy 2761 CTACAGCTCTCCCTGAGAGCGGAGTCAAGTTGAACAGCCCTGACGAGCTGACGGGCG 2820
Db 2731 CTACAGCTCTCCCTGAGAGCGGAGTCAAGTTGAACAGCCCTGACGAGCTGACGGGCG 2790
Qy 2821 GCCCTGAGAGCCAGCTTCCCTGAGAGGAGGAGAGAGAGTGTGAGAGACCAAGAGAGCT 2880
Db 2791 GCCCTGAGAGCCAGCTTCCCTGAGAGGAGGAGAGAGAGTGTGAGAGACCAAGAGAGCT 2850
Qy 2881 GAAGAGAGATCCAGGACCTCAGGACCATAGAGATGAATCCAGGCAATTTGAGCT 2940
Db 2851 GAAGAGAGATCCAGGACCTCAGGACCATAGAGATGAATCCAGGCAATTTGAGCT 2910
Qy 2941 CTTGTAACAGCTGTAAGTATCAAGACCTGAGAGGAGGCTAAACAGCTGACCGAG 3000
Db 2911 CTTGTAACAGCTGTAAGTATCAAGACCTGAGAGGAGGCTAAACAGCTGACCGAG 2970
Qy 3001 GACACGCTGAACTCAACAAACCAAACTTTACTTGTCCAAACAATCTGATGAGCTTCT 3060
Db 2971 GACACGCTGAACTCAACAAACCAAACTTTACTTGTCCAAACAATCTGATGAGCTTCT 3030
Qy 3061 GGCGCCAAGCAGAGATTGTAACAACGAGAGTGAAGTGAACCATCTCCGCGGAGATC 3120
Db 3031 GGCGCCAAGCAGAGATTGTAACAACGAGAGTGAAGTGAACCATCTCCGCGGAGATC 3090
Qy 3121 ACGGAAAGAGAGATGAGCTTACCAAGCAGAGCAAAACAGATGAGGCTGGAAGACAG 3180
Db 3091 ACGGAAAGAGAGATGAGCTTACCAAGCAGAGCAAAACAGATGAGGCTGGAAGACAG 3150
Qy 3181 TGACCATGCTGAGAGAAACAGGTCATGATTTGAGAGCCCTAAACGATGAGCTGTAAG 3240
Db 3151 TGACCATGCTGAGAGAAACAGGTCATGATTTGAGAGCCCTAAACGATGAGCTGTAAG 3210
Qy 3241 AAAGAGCGGAGTGGAGGCTGGAAGAGCGTCTGAGGATGAGAAATCCAGTTTGA 3300
Db 3211 AAAGAGCGGAGTGGAGGCTGGAAGAGCGTCTGAGGATGAGAAATCCAGTTTGA 3270
Qy 3301 TGTGCGGTTTGAAGCTGAGAGATGCTGGAACCGAAGAAACAGAGCGGCGAGAGCC 3360
Db 3271 TGTGCGGTTTGAAGCTGAGAGATGCTGGAACCGAAGAAACAGAGCGGCGAGAGCC 3330
Qy 3361 GATCAGCGGATCAGCAGTCTGCGCAGGTTGTGAGAGCTGAGTGAAGAGCAAGGCT 3420
Db 3331 GATCAGCGGATCAGCAGTCTGCGCAGGTTGTGAGAGCTGAGTGAAGAGCAAGGCT 3390
Qy 3421 GAGATCTGCTCTGAGAGGCTCTCAAGAGCAGAACTGAAAGGCCGAGGCTCTCT 3480
Db 3391 GAGATCTGCTCTGAGAGGCTCTCAAGAGCAGAACTGAAAGGCCGAGGCTCTCT 3450
Qy 3481 GACAAAGCTCAATGACTGAGAGAAAGCATGCTATGCTTGAATGATGCCCCGAAGCTTA 3540
Db 3451 GACAAAGCTCAATGACTGAGAGAAAGCATGCTATGCTTGAATGATGCCCCGAAGCTTA 3510
Qy 3541 CAGCAGAGCTGAGACTGAACGAGAGCTCAACAGAGGCTTGTGAGAGCAGAGCAAA 3600
Db 3511 CAGCAGAGCTGAGACTGAACGAGAGCTCAACAGAGGCTTGTGAGAGCAGAGCAAA 3570
Qy 3601 TTAGAGAGCAGATGAGCCTGAGAAATCAATTTCCGCTGACTCAAGGACTGCA 3660
Db 3571 TTAGAGAGCAGATGAGCCTGAGAAATCAATTTCCGCTGACTCAAGGACTGCA 3630
Qy 3661 GAGGCTTAGATCGGCTGATCTAGTGAAGACAGAAAGATGACTTGAAGTATGAGCTG 3720
Db 3631 GAGGCTTAGATCGGCTGATCTAGTGAAGACAGAAAGATGACTTGAAGTATGAGCTG 3690
Qy 3721 GAAAAATTCAGGTTCTTATTTCTATGAAGAGTGAAGAGGACATATTTCTCA 3780
Db 3691 GAAAAATTCAGGTTCTTATTTCTATGAAGAGTGAAGAGGACATATTTCTCA 3750
Qy 3781 CAAACCAATCATGATTTTCTGAAAGCAAAATGAGCAAACTGCTAAAGAAAGAA-- 3838
Db 3751 CAAACCAATCATGATTTTCTGAAAGCAAAATGAGCAAACTGCTAAAGAAAGAAAG 3810
```

```
QY 3839 -----AGTTTCTCTGCACTAC 3855
Db 3811 GGTATTATTAGTCGACGGAAAGAGACCCTGCTTACCAACACAGGTTCTCTGCACTAC 3870
QY 3856 AATGAGCTGAGCTGGGCTTGAGAGAGAAAGCTGCTGTGTGAGGCTAGAGGAAGCC 3915
Db 3871 AATGAGCTGAGCTGGGCTTGAGAGAGAAAGCTGCTGTGTGAGGCTAGAGGAAGCC 3930
QY 3916 CTTTCAAGAACCCGCACTGAGCTCCGGTCCGCCGGAGAGAGCTGGCCACCGCAAGCA 3975
Db 3931 CTTTCAAGAACCCGCACTGAGCTCCGGTCCGCCGGAGAGAGCTGGCCACCGCAAGCA 3990
QY 3976 ACGGACCAACCAACCCATCCACGCGCAACCGCAAGGACGACAGATGCGCATGTCCGC 4035
Db 3991 ACGGACCAACCAACCCATCCACGCGCAACCGCAAGGACGACAGATGCGCATGTCCGC 4050
QY 4036 ATGTGTGGGTTCGCGAAGACACAGCCCAATGACCTGTGCTGGCCCGCATCCAC 4095
Db 4051 ATGTGTGGGTTCGCGAAGACACAGCCCAATGACCTGTGCTGGCCCGCATCCAC 4110
QY 4096 CGCAGAAAGAGTCTTCACTCCAGAGAAATTTAGTGGGCTTTAGAGAACGATGCAC 4155
Db 4111 CGCAGAAAGAGTCTTCACTCCAGAGAAATTTAGTGGGCTTTAGAGAACGATGCAC 4170
QY 4156 CACAAATTTCTCAACCGATTCACGTAAGACTGACATGCGAGCCACAAAGTGTCTGTG 4215
Db 4171 CACAAATTTCTCAACCGATTCACGTAAGACTGACATGCGAGCCACAAAGTGTCTGTG 4230
QY 4216 TGTCTGATACCTGTGCACTTTGAGCGCCAGGCATCCAAATGTCTGCAATGTCAAGTATG 4275
Db 4231 TGTCTGATACCTGTGCACTTTGAGCGCCAGGCATCCAAATGTCTGCAATGTCAAGTATG 4290
QY 4276 TGTCAACCCCAAGTCTCAAGTCTGCAAGCCACCTGGGCTTGGCTGCAATATATGCC 4335
Db 4291 TGTCAACCCCAAGTCTCAAGTCTGCAAGCCACCTGGGCTTGGCTGCAATATATGCC 4350
QY 4336 ACACACTTCAACCGAGCTTCTGCGCTGACAAATGAACTCCCGAGTCTCCAGACCAAG 4395
Db 4351 ACACACTTCAACCGAGCTTCTGCGCTGACAAATGAACTCCCGAGTCTCCAGACCAAG 4410
QY 4396 GAGCCCAAGCAGAGCTTGTGACCTGGAAGGTGTGAAGGTGCCAGGAATTAACAACGA 4455
Db 4411 GAGCCCAAGCAGAGCTTGTGACCTGGAAGGTGTGAAGGTGCCAGGAATTAACAACGA 4470
QY 4456 GGAAGAGAGGCTGGGACAGGAAGTCAATTGTCTGTGAAGGGAATGAAGTCTCATTTAT 4515
Db 4471 GGAAGAGAGGCTGGGACAGGAAGTCAATTGTCTGTGAAGGGAATGAAGTCTCATTTAT 4530
QY 4516 GACAAATGAAACCAAGAGAGCTGACAGAGCGGCTGGAAGAAATTTAGCTGTGCTTCCC 4575
Db 4531 GACAAATGAAACCAAGAGAGCTGACAGAGCGGCTGGAAGAAATTTAGCTGTGCTTCCC 4590
QY 4576 GACGGGAGATGATATCTATTCATGTGTGCTTCCGAATCTCGCAATATACGCCAAA 4635
Db 4591 GACGGGAGATGATATCTATTCATGTGTGCTTCCGAATCTCGCAATATACGCCAAA 4650
QY 4636 GCAGATTTCCATATATCTAGAGATGGAATCTACCCGCAACCACTGTGCCCCGGG 4695
Db 4651 GCA----- 4653
QY 4696 AGAACCCTACTAGTCTAGCTCCAGCTTCCGTGACAAACAGCGCTGGTCAACGCTTA 4755
Db 4654 ----- 4653
QY 4756 GAATCAAGTGTGCGAGGTGGAGATTCTAGGAGAAAAAGCAGAGCTGATGCTTAACTG 4815
Db 4654 -----GAAAAAGCAGAGCTGATGCTTAACTG 4680
QY 4816 CTTGGAAGACTCCGCTGAGAACTGGAAGTGAATGACCTGTAGACATGAAGTGAACCTG 4875
Db 4681 CTTGGAAGACTCCGCTGAGAACTGGAAGTGAATGACCTGTAGACATGAACCTG 4740
QY 4876 CCCTTCAGTACCAAGGTGTGTGTGTGGGACCCGAGGAAGGGCTCTACGCGCTGAATGTC 4935
Db 4741 CCCTTCAGTACCAAGGTGTGTGTGTGGGACCCGAGGAAGGGCTCTACGCGCTGAATGTC 4800
QY 4936 TTGAAAAATCCCTTAACCCATATGTCGAGAAATTTGAGCAGTCTTCCAAATTTATATATC 4995
Db 4801 TTGAAAAATCCCTTAACCCATATGTCGAGAAATTTGAGCAGTCTTCCAAATTTATATATC 4860
QY 4996 AAGGACCTGAGAGAGCTACTATGATATGACAGAGAAAGCGGGCACTGTGTCTTGTGAC 5055
Db 4861 AAGGACCTGAGAGAGCTACTATGATGACAGAGAAAGCGGGCACTGTGTCTTGTGAC 4920
QY 5056 GTGAGAGAAATGAAACAGTCTCTGCGCAATGCCACTGCTGCGCCAGCCGACATCTCA 5115
Db 4921 GTGAGAGAAATGAAACAGTCTCTGCGCAATGCCACTGCTGCGCCAGCCGACATCTCA 4980
QY 5116 CCCAACAATTTTGAAGGTGTCAAGGGGTGCACTGTTTGGGGAGGCAAGATTGAAGAC 5175
Db 4981 CCCAACAATTTTGAAGGTGTCAAGGGGTGCACTGTTTGGGGAGGCAAGATTGAAGAC 5040
QY 5176 GGGCTCTGCACTGTGAGGCAATGCCAGCAAGTCTCATTTCTCCGCTACAAAGAAAC 5235
Db 5041 GGGCTCTGCACTGTGAGGCAATGCCAGCAAGTCTCATTTCTCCGCTACAAAGAAAC 5100
QY 5236 CTGACAAATATGTCATCCGAGAAAGATAGACCTCAGAGCCCTGCACTGTATCCAC 5295
Db 5101 CTGACAAATATGTCATCCGAGAAAGATAGACCTCAGAGCCCTGCACTGTATCCAC 5160
QY 5296 TTCAACCAATTAAGTATCTCTCAATTGGAACCAATTAATTTCAACCAATTTCAAGAACAG 5355
Db 5161 TTCAACCAATTAAGTATCTCTCAATTGGAACCAATTAATTTCAACCAATTTCAAGAACAG 5220
QY 5356 TACAGCTGAGAGAAATCTGATTAAGAAATGACCATTCCTTGGCACTGCTGTGTTGCC 5415
Db 5221 TACAGCTGAGAGAAATCTGATTAAGAAATGACCATTCCTTGGCACTGCTGTGTTGCC 5280
QY 5416 GCCTCTTCCAAAGCTTCCCTGTCTCAATGTGCAAGGTGAACAGCGAGGACGAGAG 5475
Db 5281 GCCTCTTCCAAAGCTTCCCTGTCTCAATGTGCAAGGTGAACAGCGAGGACGAGAG 5340
QY 5476 GAGTACTGTGTGTTCACAGAAATTTGAGGTGTGTGTGATTTCTTACGGAAGGCTTATG 5535
Db 5341 GAGTACTGTGTGTTCACAGAAATTTGAGGTGTGTGTGATTTCTTACGGAAGGCTTATG 5400
QY 5536 CGCAGACAGATCTCAAGGTGAGTGTGCTTACCTTTGGCTTTGGCTACAGAAACCTTAT 5595
Db 5401 CGCAGACAGATCTCAAGGTGAGTGTGCTTACCTTTGGCTTTGGCTACAGAAACCTTAT 5460
QY 5596 CTGTTTGTGACCACTTCAACTCACTGAAAGTATGAGATCCAGGACGCTCTCAGCA 5655
Db 5461 CTGTTTGTGACCACTTCAACTCACTGAAAGTATGAGATCCAGGACGCTCTCAGCA 5520
QY 5656 GGGACCCCTGCGCGAGGTATCTGAGATATCCGAAACCCGGGCTTACCTGGGCTTGGCTTAT 5715
Db 5521 GGGACCCCTGCGCGAGGTATCTGAGATATCCGAAACCCGGGCTTACCTGGGCTTGGCTTAT 5580
QY 5716 TCCCAAGAGAGATTTACTTGGGCTCTCATACAGAGTAAATTAAGGTCATTTGCTGC 5775
Db 5581 TCCCAAGAGAGATTTACTTGGGCTCTCATACAGAGTAAATTAAGGTCATTTGCTGC 5640
QY 5776 AAGGAGAACTGTGAGAGAGTCCGAGCTGAACCAACCAACCGGGGCTGCTCACTCCGC 5835
Db 5641 AAGGAGAACTGTGAGAGAGTCCGAGCTGAACCAACCAACCGGGGCTGCTCACTCCGC 5700
QY 5836 AGGAGCCCAACCAAGAGGCTCCACCTGATCAAGAGCATCACAGAGCTGTGGCC 5895
Db 5701 AGGAGCCCAACCAAGAGGCTCCACCTGATCAAGAGCATCACAGAGCTGTGGCC 5760
QY 5896 TCCAGCCAGGCGCGCGAAGGCGCCAGGACCCGAGAGGCAAGCAACCCACCGC 5955
Db 5761 TCCAGCCAGGCGCGCGAAGGCGCCAGGACCCGAGAGGCAAGCAACCCACCGC 5820
QY 5956 TACCGAGAGGGGCGAGCCGAGCTGCGCAGGAGCAAGTCTTCTGGCGCGCCCTGGAGCGA 6015
```


|||||
Db 5821 TACCGAGGGGCGGACCGAGCTCCGACGGAAGAAGTCTCTGGCCCGCCCTTGGAGCGA 5880
QY 6016 GAGAGTCCCCCGGCGATTACTCAGACCGCGGAGAGCGGTCCCCCGGAGGCTTTT 6075
Db 5881 GAGAGTCCCCCGGCGATTACTCAGACCGCGGAGAGCGGTCCCCCGGAGGCTTTT 5940
QY 6076 GAGAGTCCCCCGGCGATTACTCAGACCGCGGAGAGCGGTCCCCCGGAGGCTTTT 6135
Db 5941 GAGAGTCCCCCGGCGATTACTCAGACCGCGGAGAGCGGTCCCCCGGAGGCTTTT 6000
QY 6136 AACAGGCTGTGAGCAGCTTTC 6158
Db 6001 AACAGGAGGAGGCGGAGCTGC 6023

RESULT 15
AF086824
LOCUS
DEFINITION Mus musculus rho/rac-interacting citron kinase (Crik) mRNA,
complete cds.
ACCESSION AF086824
VERSION AF086824.1 GI:3599508
KEYWORDS
SOURCE Mus musculus (house mouse)
ORGANISM

REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 6954)
Di Cunto, F., Calautti, E., Hsiao, J., Ong, L., Topley, G., Turco, E. and
Dotto, G.P.
Citation rho-interacting kinase, a novel tissue-specific ser/thr
kinase encompassing the Rho-Rac-binding protein Citron
J. Biol. Chem. 273 (45), 29706-29711 (1998)
9909084
9792683
2 (bases 1 to 6954)
Di Cunto, F., Calautti, E., Hsiao, J., Ong, L., Topley, G., Turco, E. and
Dotto, G.P.
Direct Submission
Submitted (25-AUG-1998) Biology, University of Torino, Via Santena
5 bis, Torino, TO 10126, Italy
Location/Qualifiers
1. 6954
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/tissue_type="spleen"
1. 6954
/gene="Crik"
772. 6939
/feature="Crik"
/function="protein kinase"
/note="Crik: rho/rac binding protein; alternatively
spliced form of citron encoded by Genbank Accession Number
AF03218; similar to Mus musculus citron-K encoded by
Genbank Accession Number AF070066; similar to Rat
norvegicus citron-K encoded by Genbank Accession Number
AF070065"
/codon_start=1
/product="rho/rac-interacting citron kinase"
/protein_id="AAC72823.1"
/db_xref="GI:3599509"
/translation="MLKPYGVNPNPEASASPIASRLNLFQKPLMTTOOMS
ALSREGMDLALFEECSQPALMKHVSFVQYSDTIALRELOSPARFEVSL
VCGHFAVOYREKATGVDVAMKIMKKALAOBOVSFEERNTLSRSTSPMT
OYAFODKNNLYLMEYVGGDELSTLNKREDLDSMTQFYLAELILAVSHVQNGYV
HSDIPENILIDRTSEIKLVDFGSAKANSKVDKLPITGPDYMAEVLTVNEDRR
GTYGLDCMWSVGVAVENVYGTPTFTGTSARTNNIMNFQRLKFPDDPVSSEL
DLGLICVQERLKEBGLCHPFPARDMMNIRNSPPFVTLKSDAKTSDEK
NSMAFICVAPAPLAFGEELPFVFSYKALGYKNSVSVSLSDPSMSMKL
LISKELDSDOKKHMEQEMRLRVRSEVAVTSQREVKASRTSLLEQAT
YTBSGLKRSLEQARMVSGQBDKALQDLHDIRQSKLDEIKQETQAOVEERML

ORIGIN

Query Match 82.8%; Score 5105.2; DB 10; Length 6954;
Best Local Similarity 89.6%; Pred. No. 0;
Matches 5533; Conservative 0; Mismatches 628; Indels 11; Gaps 4;

MANLEEDVYSARRSDLYSESLAESFLAEFRKANECOHKIMAKDQKPEVGYIS
KLEKINAEQOLKTELQERKLEKAVKSTATTELLQNRQKAESELEKLNHREDS
EGIKKLVAEERHLENKVRKLEMTERRNKLDQIKTSROIQMDKILELEK
HREAOVSQAOHLKQKQYBEKIKVLNDKQDLADSELENNQAREEAEHAG
KLSEBOKAMINAMDSKIRSEORIVLESENKQANSLPTONMKAQOEMISELRO
KPYLETQACKLEAONKLEOLEKISHOHDSRSLLELETRLEYSLEBEKLEK
KQLETOQLSIOERESQUTALQARALASQLOKATLELETTREABEIOALTANDE
LOKFPALNNSCTVITDLEQLQNLQEDNAELNOMFYSLSKQDEASANDTVOLES
EVDHLREIETERMOLTSQKQTEALKTCTMLEEVLDELLEANDLEKEREMWR
SVLGDSEQFECRELOMDLEKOSRRADQRTESRQVLEAVENHAEILIALQ
ALKEOKLKAESLSLDKLEKGMENNAISLOQLLETERLEKORLEEOALQOOM
DLOKNIPLTQSLQALDLKTERSDLEVOLEINOVIVSHEKVRKEGTSOOT
LIDPLQAKNDQPAKKVPLQYNEKLALEKRAKAELEBALQKRIELRAREBAH
HRAKDHPHSTPATRQOIMASAVRSPEHQSASLAPBSRRESSTEBESRR
LKERMHNTPHRPNVGLNRRATCAVCLDTVHGRQASCLEQVWCHKEKSCPLAT
CGLPAYATHFTEAFCDKQNSPGLQSKESGSLHEGMMKVPKRNKRQOQKDRKYI
VLEGSVLIYDNEAREAGORPVEEFELCPDGVSHGAVASELANTKADVPYLIK
MESHPHTQMPGRTLYILASPFDQKRWYALASVAVGRVREKAEADKILGNLIL
KLEGDRDLNMCCTLPSPDOVYVTEBESHYALANVLKNSLTHIPGICAVQIYIYIKLE
KILMTAGEBERALCLDVKVKOSLAQSHLPADQVSPNFEAVGCHLAAGKILENSL
CICAMPSPKRVILIRYNDNLSKYCIRKEIETSPSCIHPTNYSILGINKFYEIDMKO
YTLDEFLDKNDHSLAPAVPASSNSFPVSIQVANSAGQREYLLCFHEGVEFDSYGR
RSRTDLKWSRLPLAVAREPLYTFHFSLEVIEIQASSLSGPARAVLEIPNRYL
GPAISGATYLSVYODKLRVYCKGNLYKESGTEQHRVPSTSRSPNKRGPTEVH
ITKRAVSSPAPEGSHPEPSTPHRYDRREGTELRKDSFPERPLEKESPERMLST
KRERSFGRLEFDSRGRPLPAGAVRTPLSQVNRVMDQSV"

QY 1 ATGTGAACTCAATATGAGCGCGGAATCTTTGAGTCTGCTGCTGAAACCCATT 60
Db 772 ATGTGAACTCAATATGAGTGTGCGGAACCCGCGAGCCAGTGCCTCGAGCCATT 831
QY 61 GCCAGCCGGGCTCCAGGCTGAATCTGTTCTTCAGGGGAAACACCTTTATGACTCA 120
Db 832 GCCAGTCCGGGCTCCAGGCTGAATCTTCTTCAGGGGAAACCGCCCTCATGACTCAA 891
QY 121 CAGCAGATGCTCTCTCTTCCCGGAGAGGATTAAGATGACCTTTGTTCTTGGAA 180
Db 892 CAGCAGATGCTCTCTTCCCGGAGAGGATTAAGATGACCTTTGTTCTTGGAA 951
QY 181 GAATGACAGTCAAGCTCTCTGTAAGAAATTAAAGCATGAGCACTTTGTCCGGAAGTAT 240
Db 952 GAGTGAAGCAACCGCCCTGATGAAGATGAAGCACTGAGCAGCTTTGTCCAGAAAGTAT 1011
QY 241 TCCGACACCATAGCTGAGTTACAGAGCTCCAGCTTGGCAAAGACCTTGAAGTCA 300
Db 1012 TCCGACACCATAGCTGAGTTACAGAGCTCCAGCTTGGCAAAGACCTTGAAGTCA 1071
QY 301 AGCTTGTAGGTGTGTCATCTTCTGAAGGAGGTGTGAAGAGAAAGCAACCGGG 360
Db 1072 AGCTTGTAGGTGTGTCATCTTCTGAAGGAGGTGTGAAGAGAAAGCAACCGGG 1131
QY 361 GACATCTATGCTATGAAAGTGAAGAAAGGCTTTATTTGCGCCAGAGCAGGCTTCA 420
Db 1132 GACGCTATGCGATGAATAATATGAAGAAAGGCTTTGCGCCAGAGCAGGCTTCA 1191
QY 421 TTTTGTGAGAAAGCGGACATATTATCTGAAAGCAAGCCCTGATCCCCCAATTA 480
Db 1192 TTTTGTGAGAAAGCGGACATATTATCTGAAAGCAAGCCCTGATCCCCCAATTA 1251
QY 481 CAGTATGCTTTCAGGACAAAATACTTATCTGTCGTCATGGAATACACGCTGGAAGG 540
Db 1252 CAGTATGCTTTCAGGACAAAATACTTATCTGTCGTCATGGAATACACGCTGGAAGG 1311
QY 541 GACTTGTCTGACTTTGAATAGATATGAGACCAAGTATGAAACCTGATCAAGTTT 600
Db 1312 GATTTCCTGTGCTTTCGAAAGATACAGAGACCAATTAGATGAGAGATATCAAGTTT 1371
QY 601 TACCTAGCTGAGCTGATTTGCTGTTCAACAGCTTATCTGATGGAATACGATGATCA 660

Db 1372 TACCTTGCTGAGTGATTTTGGCTGTCCACAGCGTGCACAGATGGATTTGTGCTCGA 1431
Qy 661 GACATCAAGCGCTGAGAACATTCTGTTGACCGCACAGACACATCAAGCTGTGGATTTT 720
Db 1432 GACATCAAGCGCTGAGAACATTCTGTTGACCGCACAGACATCAAGCTGTGGATTTT 1491
Qy 721 GGA TCTGCCGCGAAAAATGAAATTCMAACAGATGGTGAATGCCAACTCCGATTTGGGACC 780
Db 1492 GGATTCAGCGGCTTAAGATGAATTCAAATTA ---GGTGAATGCCAAACTCCCATTTGGGACC 1548
Qy 781 CCGAATTCATCTGGCTCTCTGAAGTCTGACTGTGTATGAACGGGGATGGAAAAAGCACTTAC 840
Db 1549 CCGGAATTCATCTGGCTCTCGGAAGTGTGACCGTATGAAACGAGACCGAAAGGGCACATAC 1608
Qy 841 GGGCTGAGACTGTACTGCTGTGTGAGTGGGGGTGATTGCTTATGAGATGATTTATGGGAGA 900
Db 1609 GGGCTGAGACTGTGACTGTGTGAGTGGGGGTGATTGCTTATGAGATGATTTATGGGAG 1668
Qy 901 TCCCTCTTCCAGAGGGGAACTCTGCGCAACCTTCATTAACATTAATGAATTTCCAGCGG 960
Db 1669 ACCCAATTCAGAGGGGAACTCTCGGCCGAGACCTTCAAACAATCATGAATTTCCAGCGG 1728
Qy 961 TTTTGAATTTTCCAGATGACCCCAAGTGAAGTGAAGTCTTTGATCTGATTCAAAGC 1020
Db 1729 TTTTGAAGTTCCAGATGACCCCAAGTGAAGTGAAGTCTTTGATCTGATTCAGAGT 1788
Qy 1021 TTGTGTGGCGCGCAGAAAGAGAGAGTGAAGTTGAAGTCTTTGCGCACTCCCTTCTTC 1080
Db 1789 CTCTGTGTCCAGAAAGAGAGCTGAAGTTCAGGGGTCTCTGTGCCACCTTCTTCTT 1848
Qy 1081 TCTTAAATTTGACTGSAACAATTCGTAACTCTCTCCCTTCGTTCCCACTTCAG 1140
Db 1849 GCCAGAAAGGACTGGAACCAATCCGTAATCTCTCCCTTCGTTCCCACTTCAG 1308
Qy 1141 TCTGAGATGACACTTCCAAATTTTGTATGAACAGAGAAAGATTCGTGGG- TTTTATCTTC 1199
Db 1909 TCTGAGATGACACTTCCAAATTTTGTATGAACAGAGAAAGATTCGTGGGCTTCATCTTC 1968
Qy 1200 TCCGTGCGAGCTGAGCCCTCAGGCTTCGTGGGGAAGAACTCCGTTTGTGGGTTTC 1259
Db 1969 TGTGTGCAAGCTAGAGCCCTC- GCGTTCTCAGGCGAAAGCTCGCGTTTGTGGATTTTC 2027
Qy 1260 GTACAGCAAGGCACTGGGGAATTCGTGGTATGATCTGAGTCTGTTGTGCGGCTGAGCTC 1319
Db 2028 GTACAGCAAGGCACTGGGGAATTCGTGGTATGATCTGAGTCTGTTGTGCGGCTGAGCTC 2087
Qy 1320 CCCTGCCAAGACTAGCTCCATGGAAGAAAGAACTTCTCATCAAAAGCAAAAGCTCCAAAG 1379
Db 2088 CCCTGCCAAGGTTAGCTCCATGGAAGAAAGAACTTCTCATCAAAAGCAAAAGCTCCAAAG 2147
Qy 1380 CTCTCAGCAAGATGTCTCAAAAGTGAAGCAGAAATACCCGCTTAACTGAGAGTTC 1439
Db 2148 CTCTCAGCAAGATGTCTCAAAAGTGAAGCAGAAATACCCGCTTAACTGAGAGTTC 2207
Qy 1440 AGAGGTGAGAGCTGTGCTTATGTCAGAAAGAGTGAAGCTGAAGGCTCTGAGACTCAG 1499
Db 2208 AGAGGTGAGAGCTGTGCTTATGTCAGAAAGAGTGAAGGCTCTGAGACTCAGAG 2267
Qy 1500 ATCCCTCTTGGAGCAGGACCTTGCTACTTACATCAAGAAATGCAAGTACGTTAAAGCGAG 1559
Db 2268 ATCCCTCTTGGAGCAGGACCTTGCTACTTACATCAAGAAATGCAAGTACGTTAAAGCGAG 2327
Qy 1560 TTTTGAAGCAAGCAGATGAGAGTGTCCAGAGAGATGACAAAGCACTGAGCTTCTCCA 1619
Db 2328 TTTTGAAGCAAGCAGGAGTGAAGTGTCCAGAGAGATGACAAAGCTTGAAGCTTCTCCA 2387
Qy 1620 TGAATATCAGAGAGCAGGACCGGAAGCTCAAGAAATCAAAAGAGAGGATACAGAGCTCA 1679
Db 2388 CGAGATCCGAGAGAGAGCCGGAAGCTTCAGAGATCAAGAGAGAGATACAGAGCTCA 2447
Qy 1680 AGTGAAGAAATGAGTTGATGATGAATCAAGTTGAAGAGATCTTGTCTCAGCAAGAG 1739

Db 2448 GGTGAGAGATGAGGCTGATGATGAATCAGCTGGAAGAAAGACTGTGTACGCCGAG 2507
Qy 1740 ACCGAGTGAATCTCTAAGAAATCTGAGCTGAGAGAGTCTCGGCTTGTCTGTAAGAAATCAA 1799
Db 2508 ACCGAGTGAATCTCTAAGAGTCTGAGAGAGAGTCTCGGCTTGTCTGTAAGAAATCAA 2567
Qy 1800 GCGGAAAGCAGCAATGTCAAGATTAATCTTTGAAGGCTTAAGATCAAGGAAAGCTTGA 1859
Db 2568 GCGGAAAGCAGCAATGTCAAGCAAACTGAATGAAGGCTTAAGAGACCAAGGAAAGCTTGA 2627
Qy 1860 AGTGGAGAAATATGCAAACTGGAAGAAATCAATGCTGAGCAGACAGCTCAAAATTCAGGA 1919
Db 2628 AGTGGAGAAATATTCAAACTGGAAGAAATCAATGCTGAGCAGACAGCTGAATTCAGGA 2687
Qy 1920 GCTCCAAAGAAACTGGAAGAGCTGTAAAGCCAGACAGAGGCGCACAGAGTGTCTGA 1979
Db 2688 GCTCCAGAGAACTGGAAGAGCTGTAAAGCCAGACAGAGGCGCACAGAGGCTGTGCA 2747
Qy 1980 GAATATCCGCAAGGCAAGAGAGCGAGCCGAGAGAGCTGGAAGAGCTGCAGAACCGAGA 2039
Db 2748 GAATATCCGCAAGGCAAGAGAGCGAGCCGAGAGCTGGAAGAGCTGCAGAACCGAGA 2807
Qy 2040 GGAATCTTCTGAAGGCAATGCAAGAAAGAGCTGTGGAAGCTGGAAGAGCGGCAATCTCT 2099
Db 2808 AGACTCTCTCGAGAGGCAATCAAAAGAGAGCTGTGGAAGCGAGAGAGCGGCAATCTCT 2867
Qy 2100 GGAGAAACAAGGTAAAGACTAGAGCAATGAGAGCTGAGAGAAAGAGAGAGTGA 2159
Db 2868 GGAGAAACAAGGTAAAGAGACTAGAGCAATGAGAGCTGAGAGAAAGAGAGTGA 2227
Qy 2160 CATCCAGCAAAATCCCAAGATCCAGAGAGAGAGTGTGATTAATTTGTGAGCTCGAAG 2219
Db 2928 CATCCAGCAAAATCCCAAGATCCAGAGAGAGAGTGTGATTAATTTGTGAGCTCGAAG 2287
Qy 2220 GAAACATCGGAGAGCCCAAGTCTCAGCCCAAGCACTGAAGTCACTGTAACAGAAAGA 2279
Db 2988 GAAACATCGGAGAGCTCAGGTCTCAGCTCAACATCTGAAGATCACTTGAACAGAAAGA 3047
Qy 2280 GCAGCACTATGAGAGAAAGATTAAGTGTGAGCAATCAGATTAAGAAAGCACTGAGCTGA 2339
Db 3048 ACAGCACTATGAGAGAAAGATTAAGTGTGAGCAATCAGATTAAGAAAGCACTGAGCTGA 3107
Qy 2340 CAAAGAGCACTGAGAGCAATGATGCAAGAGACAGAGAGAGGCGCATGAGAAAGGCA 2399
Db 3108 CAAAGAGAGCTGAGAGCAATGATGCAAGAGACAGAGAGAGGCGCATGAGAAAGGCA 3167
Qy 2400 AATTTCAAGGAAACAGAGGCAATGATCAATGCTATGAGATTTCCAAATCAATCCCTGGA 2459
Db 3168 GATCTTCAAGGAGCAAGAGGCGATGATCAACGAGATGGAATTCMAAGATCCGATCCCTGGA 3227
Qy 2460 ACAGAGGATTTGGAAGCTGTCTGAGGCAATTAATCTTGAGCAAAATGAGAGCTTTTAC 2519
Db 3228 GCAAGAGATGTGAGAGCTGTGAGAGGCAAGAGCTTGCGCAAAAGAGAGCTTTTAC 3287
Qy 2520 CCAAGAGCAATGAGAGGCGCCCAAGAGAGATGATTTCTGAATCTGAGCAACAGAAATTTTA 2579
Db 3288 CCAAGAGCAATGAGAGGCGCCCAAGAGAGATGATCTGAGAACTCAGGCAAGAGAAATTTTA 3347
Qy 2580 CTTGAGACACAGGCTGGGAAGTTGAGAGGCGCCAGAACCAAAATCTGAGAGAGAGCTGGA 2639
Db 3348 CTTGAGAGACAGGCGCGGAAGGCTGAGAGGCGCCAGAACCAAAATCTGAGAGAGAGCTGGA 3407
Qy 2640 GAAGATCAGGCAACAGACCAAGTGAAGAAATGCGGCTGTGGAATCTGAGAGCAAGATT 2699
Db 3408 GAAGATCAGGCAACAGAGATCAGTGAAGAAATGCGGCTGTGGAATCTGAGAGCAAGATT 3467
Qy 2700 GCGGAGAGTCAAGTCTAGAGCAGAGAGCAGAAATCTGAGAGCTCAAGCGCAGCTCAGAGA 2759
Db 3468 GAGGAGAGTCAAGCTGAGAGCAGAGAGCAGAAAGCTGAGAGCTGAAGGCGAGCTCAGAGA 3527
Qy 2760 GCTTACAGCTTCTCTGAGAGAGCGGCAAGTCAAGTTGACAGGCTTGAAGGCTGACAGGCG 2819
Db 3528 GCTTACAGCTTCTCTGAGAGAGCGGCAAGTCCAGCTGACGCGCTTGAAGGCGCGGCG 3587

QY	2820	GGCCCTTGGAGACGCGCTTTCGCCAGCGGAAAGACAGAGCTGGAAAGACCAACAGCAGAAAGC	2879
Db	3588	AGGCGTGGAGAGCCAGCTCCGCCAGGCGAAAGACAGAGCTGGAGAGCAACCCCGGAAGC	3647
QY	2880	TGAAGAGAGATCCAGGACCTCAACGGCACTAGAGATGAAATCCAGCCGAATTTGATGC	2939
Db	3648	GGAGAGAGAGATCCAGGCGCTCAACGGCACTCCGATGAGATCCAGCGCAAAATTCATGCG	3707
QY	2940	TCTTCGTAAAGCTGTACTGTAACTCAAGACTTGGAGAGCAGCTAAACAGCTGACCGA	2999
Db	3708	CTTCGCAACAGCTGACCGGTCTATCAACCGACTGGAGAGCAGCTGACAGCTCACCGA	3767
QY	3000	GGACAACGCTGAATCAACAACCAAACTTCTACTGTCCAAACAACTGATGAGGCTTC	3059
Db	3768	GGACAACGCGAGACTCAACAACCAAACTTCTACTGTCCAAACAACTGATGAGGCTTC	3827
QY	3060	TGCGCGCAACGAGAGATTTGTACAACTGCGAAGTGAAGTGAACATCTCGCGCGGAGAT	3119
Db	3828	CGGGGCGCAATGACGAGATTTGTGACCTGCGAGTGAAGTGAACATCTCGCGGTGAGAT	3887
QY	3120	CACGGAACGAGAGATSCAGCTTACAGCGAAGAAAGCAAGTGAAGGCTCTGAAAGCCAC	3179
Db	3888	CACGAGACGGGAGATCAGCTCAACGACCAAGAAAGCAACGATGAGGCTCTGAAAGCCAG	3947
QY	3180	GTGSCACCATGCTGGAGAGAAACAAGTCAATGATTTGGAGGCGCTTAACAGATGAGCTGTAGA	3239
Db	3948	ATGCAACCATGTTGGAGAGACAGAGTCTGTGACCTGGAGGCGCTTGAAAGAGAGAGCTGTGGA	4007
QY	3240	AAAAAGCCGAGTGGAGAGCGCTGAGAGAGCGTCTGGGTGATGAGAAATTCGCAAGTTTA	3299
Db	4008	GAAAGAGCGCCAGTGGAGAGCGCTGGGAGAGCGTCTTGGCGACGAGAAATCCCAATTGCA	4067
QY	3300	GTGTGCGGTTCCAGAGCTGCAAGAGATGCTTGAACACCGAAGAACAGAGCAGGCGAGAGC	3359
Db	4068	GTGTGCAAGTTCCAGAGCTCAAGAGATGCTTGAACACCGAAGAACAGAGCAGGCGCTTGAAGGCG	4127
QY	3360	CGATCAGCGGATCACCGAGTCTGCCAGGTGGAGGCTGGGAGTGAAGAGAGACACAGGCG	3419
Db	4128	CGATCAGCGGATCACCGAGTCCGCGCGAGGTGGAGTGGCGGTGAAGAGAACACAGGCG	4187
QY	3420	TSAGATTCTCGCTTTCGACAGAGCTTCTCAAGAGCAGAAAGCTGAAGGCCGAGAGCGCTCTC	3479
Db	4188	CGAGATTCTTGTCTTTCGACAGAGCTTCTCAAGAGCAGAAAGCTCAAGCGGAGAGCGCTGTC	4247
QY	3480	TCACAAGCTCAATGACCTTGGAGAGAGAGATGCTATGCTTGAATGAAATGCCGAAGCTT	3539
Db	4248	GGACAAGCTCAACGACCTGGAGAGAGAGACCGCACTGCTGAGATGAAGCGCCCGAGGCTT	4307
QY	3540	ACAGAGAAAGCTGAGAGACTGAACAGAGAGCTCAAAACAGAGGCTTCTGAGAGCAAGCCAA	3599
Db	4308	ACAGAGAAAGCTGAGAGACGAGACCGGAGCTCAAAACAGAGGCTTCTGAGAGCAAGCCAA	4367
QY	3600	ATTACAGACAGATGAGCTTCGACAGAAATATCAATTTCCGTCTGACTCAAGACTGCA	3659
Db	4368	ATTACAGACAGCTGAGATGACCTTCGACAGAAACCAATTTCAAGACTGACGCAAGGCGTCA	4427
QY	3660	AGAGCTCTAGATTCGGGCTGATCTACTGAAACACAAAGAAAGTGAACCTTGGAGATCACT	3719
Db	4428	GGAGGCGCTGGACCGGCGGAGTCTGTAAGACACAAAGAGAGCGACCTGGAGATCCACACT	4487
QY	3720	GGAAAAATTCAGGTTCTCTATTCATATGAAAAAGGTGAAGGCACTAATTTCTCA	3779
Db	4488	GGAAAAATTCAGGTTCTCTACTCTCAAGAGAAATGAAATGAAAGGCACTAATTTCTCA	4547
QY	3780	ACAAACCAAACTCATTTATTTTCTGCAAGCCAAATGAGCAACCTGTGTAAGAAAAA	3839
Db	4548	GCAAAACCAAACTCATTTGATTTCTGCAAGCCAAATGAGCAACCTGTGTAAGAAAAA	4607
QY	3840	GGTTCCTGTGAGTACATGAGCTGAGAGCTGGCCCTGGAAGAGAGAAAGCTCGCTGCGC	3899
Db	4608	GGTTCCTGTGAGTACAAATGAGCTGAGAGCTGGCCCTGGAAGAGAGAAAGCCGATGCGC	4667

QY	3900	AAGGCTAAGAGAGGCTTTCACAAAGCCGATTCAGAGCTCCGCTCCGCGCGGAGAGAGC	3959
Db	4668	GGAGCTGAGAGAGGCTTTCAGAAAGCCGATTCGAACTTCGCGCTTCGCGCGGAGAGAGC	4727
QY	3960	TGCCCAACGCAAGAGAGAGAGCAACCCACATTCACCGCCAGCCACCGGAGGACACA	4019
Db	4728	TGCCCAACGCAAGAGAGAGAGCAACCCAGCCAGCCGCTTACCGCCAGCACTTGGAGGACACA	4787
QY	4020	GATCGCATGTCGCGCATCTGTGCGGTCCGAGAGAGACAGGCGAGTGCATGAGGCTCT	4079
Db	4788	GATCGCATGTCGTGCACTTGTGCGGTCCGAGAGACAGGCGAGTGCATGAGGCTCT	4847
QY	4080	GGCCCGGCATTCAGCGCGCAGAAAGAGTCTTCAACTCCAGAGAAATTAAGTCGCGCTCT	4139
Db	4848	TGCCCCACCATTCAGCGCGCAGAAAGAGTCTCTCAATCCAGAGAAATTCACCGGCGCTCT	4907
QY	4140	TAAAGAAAGCATGACCAACAATAATTCCTCAACCGAATTCACGTAGAGCTGAACATGCGAGC	4199
Db	4908	GAAAGAGCGCATGACCAACAATTCCTCAACCGGTTTAATGTGAGGCTGAAACATGAGAGC	4967
QY	4200	CACAAAGTGTCTGTGTCTGTGATACCGTGCACTTTGAGCGCCAGGATCCAAATGCT	4259
Db	4968	CACCAAGTGGCCGCTGTGTCTGTGATATGTGCACTTTGAGCGCCAGGATCCAAATGCT	5027
QY	4260	CGAATGTAGGTGATGTGTCAACCCCAAGTCTCAAGTCTTTCGACGCACTTCGCGCTT	4319
Db	5028	CGAATGTCAAGTATGTGTCAACCCCAATGTCTCAACCTGTTCGCTACCTGTGCGCT	5087
QY	4320	GCTGCTGAATATGACACACACTTCAACGAGGCGCTTTCGCGGTGACAAATGAACTCCCG	4379
Db	5088	GCCAGTGAATATGCCACACACTTCACTGAGGCTTTCGCGGAGCAAAATGAACTCCCG	5147
QY	4380	AGGTCTTCAGACCAAGAGCCGACAGCAAGCTTTCACCTGGAAAGGTGATGAAGTGC	4439
Db	5148	GGGCTCCAGAGCAAGAGCCTTGCACAGCTTTCACCTGGAAAGGTGATGAAGTGC	5207
QY	4440	CAGGAATTAACAAGAGAGACGAAAGCTGGAGACAGGAATGATCAATTCCTGAGGGATC	4499
Db	5208	CAGGAATTAACAAGAGAGACGAAAGCTGGAGACAGGAATGATCAATTCCTGAGGGATC	5267
QY	4500	AAAACTCTCATTTATGACATGAAAGCAGAGAGCTGACAGAGGCGGCTGAAAGATT	4559
Db	5268	AAAACTCTCATTTATGACATGAAAGCAGAGAGCTGACAGAGGCGGCTGAAAGATT	5327
QY	4560	TGAGCTGTGCTTTCGCAACGGGATGTATCTATTATGATGTCGCTGTGTCCTCCGACT	4619
Db	5328	TGAGCTGTGCTTTCGCAACGGGACGATATCTATTATGATGTCGCTGTGTCCTCCGAACT	5387
QY	4620	CGCAAAATACAGCCCAAGACAGATGCCATATCATCTGAAGATGGAATCTGACCCGACAC	4679
Db	5388	TGCAAAATACAGCCCAAGACAGATGCCATATCATCTGAAGATGGAATCTTACCCGACAC	5447
QY	4680	CACCTGTGCGCGGGAGAAACCTCTAATTTGCTTACGCTCCAGCTTCCCTGACAAACGCG	4739</

```
Db 5748 CCAAATTACATCTCAAGAGCTGGAGAGCTGCTATATACAGAGGAGAGCGGGC 5807
Qy 5040 ACTGTGCTTGTGGACGTGAAGAGTGAACAGTCCCTGGCCAGTCCCACTGCTGC 5099
Db 5808 TCTGTGCTGTGGACGTGAAGAGTGAAGCACTCCCTGGCCAGTCACTGCTGC 5867
Qy 5100 CCAAGCCGACATCTCAACCAATTTTGAAGCTGTCAAGAGCTGCCACTTTGTTGGGGC 5159
Db 5868 CCAAGCCGACATCTCAACCAATTTTGAAGCTGTCAAGAGCTGCCACTTTGTTGGGGC 5927
Qy 5160 AGGGAATTTGAAGACGGGCTTGGCATCTGTGACAGCATGCCCAGCAAGTCTGATCT 5219
Db 5928 TGGCAATGGAAGACAGCTGTGCATCTGCGCGTATGCAAGCAAGTCTGATCT 5987
Qy 5220 CCGCTTCAAGAAACCTCAGCAATATCTGCATCCGGAAAGATAGAGACTCAGAGCC 5279
Db 5988 CCGCTTCAATGACAACTCAGCAATATCTGCATCCGGAAAGTCAAGACTCAGAGCC 6047
Qy 5280 CTGAGCTGTATCCACTTCAACCAATTAAGATATCTCATTTGGAACCAATTAATTTA 5339
Db 6048 CTGAGCTGTATCCACTTCAACCAATTAAGATATCTCATTTGGAACCAATTAATTTA 6107
Qy 5340 AATGACATGAAGACATACAGCTCGAGAAATCTGATTAAGATGACCAATCTTGGC 5399
Db 6108 GATGACATGAAGACATACAGCTCGAGAAATCTGATTAAGATGACCAATCTTGGC 6167
Qy 5400 ACCGTGTGTGTTGCGGCTCTTCCAAAGCTTCCATCTGTCTCATCTGTCAGTGAAG 5459
Db 6168 ACCGTGTGTGTTGCGGCTCTTCCAAAGCTTCCATCTGTCTCATCTGTCAGTGAAG 6227
Qy 5460 CGCAGGACAGAGAGAGTACTGTGTGTTTCCAGAAATTTGAGTGTCTGATTC 5519
Db 6228 CGCGGACAGAGAGAGTACTGTGTGTTTCCAGAAATTTGAGTGTCTGATTC 6287
Qy 5520 TTACGGAAGACGTACCGCACAGATCTCAAGTGAAGTCTTACCTTGGCTTGC 5579
Db 6288 TTACGGAAGACGTACCGCACAGATCTCAAGTGAAGTCTTACCTTGGCTTGC 6347
Qy 5580 CTACAGAAACCTATCTGTTTGAACCACTTCAACTCACTCGAAGTATGATCA 5639
Db 6348 CTACAGAAACCTATCTGTTTGAACCTTCAACTCACTCGAAGTATGATCA 6407
Qy 5640 GGCAGCTCTCTCAGCAGAGACCCCTGCGCAGCTGACCTGACATCCGAACTCCGCTA 5699
Db 6408 GGCAGCTCTCTCAGCAGAGACCCCTGCGCAGCTGACCTGACATCCGAACTCCGCTA 6467
Qy 5700 CTTGGGCTCTGCAATTTCTCAGAGCATTTACTTGGCGTCTCATACAGATAATT 5759
Db 6468 CTTGGGCTCTGCAATTTCTCAGAGCATTTACTTGGCGTCTCATACAGATAATT 6527
Qy 5760 AAGGTCATTTGCTGCAAGGAAACCTCGTAAGAGTCCGGCACTGAACCAACCGGGG 5819
Db 6528 AAGGTCATTTGCTGCAAGGAAACCTCGTAAGAGTCAAGCACTGAACCAACCGGGG 6587
Qy 5820 CCGTCCACTCCCGCAGAGCCCAACAAGCAGGCCCACTCACTAACAAGCAGCAT 5879
Db 6588 GCGCTCACTCCCGCAGAGCCCAACAAGCAGGCCCACTCACTAACAAGCAGCAT 6647
Qy 5880 CACCAAGCGGTGCTTCAAGCCGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGG 5939
Db 6648 CACCAAGCGGTGCTTCAAGCCGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGG 6707
Qy 5940 AAGCAACCCCAACCGCTAC-----GCGAGGGCGGACCGAGCTGGCGGAGCAATG 5993
Db 6708 AAGCAACCCCAACCGCTAC-----GCGAGGGCGGACCGAGCTGGCGGAGCAATG 6767
Qy 5994 TCTGAGCGGCGCTTGAAGCAGAGAGTCCCGCGCGGATCTCAGCAGCGCGAGAGA 6053
Db 6768 TCTGAGCGGCGCTTGAAGCAGAGAGTCCCGCGCGGATCTCAGCAGCGCGAGAGA 6827
Qy 6054 GCGGTCCCGCGAGGCTGTTGAAGAAGCAGAGAGGCGGCGGCTGCTGCGGAGCGGT 6113
```

```
Db 6828 GCGTCCCGGAGAGTGTGTAAGACAGACAGAGGCGCGGCTGCTCAGAGCACT 6887
Qy 6114 GAGACCCGCTGTCCAGGTGAACAAGTCTGGGACCACTTTCACTATPA 6165
Db 6888 GAGACCCCACTGTCCAGGTGAACAAGTCTGGGACCACTTTCACTATPA 6939
```

Search completed: May 16, 2005, 15:06:37
Job time : 17095 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 16, 2005, 09:14:34 ; Search time 653 Seconds
(without alignments)
1548.156 Million cell updates/sec

Title: US-10-791-666-1

Perfect score: 6165
Sequence: 1 atgtcgaagtcacaatg.....gggaccgcttcagataa 6165

Scoring table: IDENTITY NUC
Gapop 10'-0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 10%
Listing first 45 summaries

Database : Issued Patents NA.*

- 1: /cgn2_6/ptodata/1/ina/5A.COMB.seq.*
- 2: /cgn2_6/ptodata/1/ina/5B.COMB.seq.*
- 3: /cgn2_6/ptodata/1/ina/6A.COMB.seq.*
- 4: /cgn2_6/ptodata/1/ina/6B.COMB.seq.*
- 5: /cgn2_6/ptodata/1/ina/PTCUS.COMB.seq.*
- 6: /cgn2_6/ptodata/1/ina/backfile1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	6165	100.0	6165	US-10-028-946-1	Sequence 1, Appli
2	5837	94.7	5877	US-10-028-946-3	Sequence 3, Appli
3	1398.2	22.7	1515	US-09-804-471A-1	Sequence 1, Appli
4	1398.2	22.7	1515	US-10-238-709-1	Sequence 1, Appli
5	958.8	15.6	2162	US-09-774-528-419	Sequence 419, App
6	756.4	12.3	1133	US-09-916-204-1	Sequence 1, Appli
7	756.4	12.3	1133	US-10-282-048-1	Sequence 1, Appli
8	258	4.2	2358	US-09-016-434-513	Sequence 513, App
9	233	3.8	2423	US-09-949-016-2640	Sequence 2640, Ap
10	227.2	3.7	2726	US-08-422-699A-12	Sequence 12, Appli
11	227.2	3.7	2726	US-08-422-706B-12	Sequence 12, Appli
12	220.4	3.6	2511	US-08-422-699A-8	Sequence 8, Appli
13	220.4	3.6	2511	US-08-422-706B-8	Sequence 8, Appli
14	220	3.6	3182	US-08-484-044-11	Sequence 11, Appli
15	205	3.3	174493	US-09-804-471A-3	Sequence 3, Appli
16	205	3.3	174493	US-10-238-709-3	Sequence 3, Appli
17	195.6	3.2	2706	US-08-630-822A-61	Sequence 61, Appli
18	195.6	3.2	2706	US-09-005-069-61	Sequence 61, Appli
19	195.6	3.2	2706	US-09-171-156A-20	Sequence 20, Appli
20	195.6	3.2	2706	US-09-004-730A-20	Sequence 20, Appli
21	195.6	3.2	2706	US-08-981-799A-20	Sequence 20, Appli
22	186	3.0	4363	US-08-685-576-5	Sequence 5, Appli
23	184.4	3.0	4848	US-09-976-594-295	Sequence 295, App
24	178.8	2.9	48763	US-09-916-204-3	Sequence 3, Appli
25	178.8	2.9	48763	US-10-282-048-3	Sequence 3, Appli
26	175.4	2.8	4065	US-09-016-434-1105	Sequence 1105, Ap
27	175.4	2.8	4739	US-08-685-871-1	Sequence 1, Appli

28	173.2	2.8	5053	2	US-08-685-576-2	Sequence 2, Appli
29	157	2.5	193	4	US-09-513-999C-20993	Sequence 20993, A
30	156.6	2.5	3323	2	US-08-422-699A-10	Sequence 10, Appli
31	156.6	2.5	3323	2	US-08-422-706B-10	Sequence 10, Appli
32	125.8	2.0	3155	3	US-09-442-100-7	Sequence 7, Appli
33	125.8	2.0	3155	4	US-08-939-106-7	Sequence 7, Appli
34	125.8	2.0	3155	4	US-09-442-102-7	Sequence 7, Appli
35	125.4	2.0	1498	3	US-09-509-902A-6	Sequence 6, Appli
36	125.4	2.0	1961	3	US-09-509-902A-15	Sequence 15, Appli
37	125.4	2.0	5276	4	US-09-233-857-2	Sequence 2, Appli
38	121	2.0	3018	2	US-08-860-150-6	Sequence 6, Appli
39	121	2.0	3018	3	US-09-338-132-6	Sequence 6, Appli
40	114.6	1.9	1935	2	US-08-878-989-11	Sequence 11, Appli
41	114.6	1.9	1935	3	US-09-272-796-11	Sequence 11, Appli
42	113	1.8	3213	3	US-09-442-100-5	Sequence 5, Appli
43	113	1.8	3213	4	US-08-939-106-5	Sequence 5, Appli
44	113	1.8	3213	4	US-09-442-102-5	Sequence 5, Appli
45	110.4	1.8	1251	4	US-09-248-796A-4400	Sequence 4400, Ap

ALIGNMENTS

RESULT 1
US-10-028-946-1
Sequence 1, Application US/10028946
Patent No. 6734009
GENERAL INFORMATION:
APPLICANT: Yu, Xuanhuan
APPLICANT: Miranda, Maricar
APPLICANT: Fridlie, Carl Johan
TITLE OF INVENTION: No. 6734009e1 Human Kinases and Polynucleotides Encoding the Same
FILE REFERENCE: LEX-0289-USA
CURRENT APPLICATION NUMBER: US/10/028, 946
CURRENT FILING DATE: 2001-12-20
PRIOR APPLICATION NUMBER: US 60/258, 335
PRIOR FILING DATE: 2000-12-27
NUMBER OF SEQ ID NOS: 4
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1
LENGTH: 6165
TYPE: DNA
ORGANISM: homo sapiens
US-10-028-946-1

Query Match 100.0%; Score 6165; DB 4; Length 6165;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 6165; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	ATGTTGAAGTTCAATATGAGCGCGAATCTTTGATGCTGTGCTGAACCCATT	60
DB	1	ATGTTGAAGTTCAATATGAGCGCGAATCTTTGATGCTGTGCTGAACCCATT	60
QY	61	GCCAGCCGGGCTCTCAGCTGAATCTTTTCCAGGGGAACCACTTTATATCTCA	120
DB	61	GCCAGCCGGGCTCTCAGCTGAATCTTTTCCAGGGGAACCACTTTATATCTCA	120
QY	121	CAGAGATGCTCTCTTCCGAGAGGATATTAGATGCTCTTTGTTCTTTGAA	180
DB	121	CAGAGATGCTCTCTTCCGAGAGGATATTAGATGCTCTTTGTTCTTTGAA	180
QY	181	GAATGAGTCAAGCTCTGATGATTAAGCACTGAGCACTTTGTCGGAAGTAT	240
DB	181	GAATGAGTCAAGCTCTGATGATTAAGCACTGAGCACTTTGTCGGAAGTAT	240
QY	241	TCCGACACCATGCTGATTAAGCACTGAGCACTTTGTCGGAAGTAT	300
DB	241	TCCGACACCATGCTGATTAAGCACTGAGCACTTTGTCGGAAGTAT	300
QY	301	AGTTTGAAGTGTGCTGCTTTGCTGAAGTGAAGTGAAGCAACCGGG	360
DB	301	AGTTTGAAGTGTGCTGCTTTGCTGAAGTGAAGTGAAGCAACCGGG	360

QY 361 GACATCTATGCTATGAAGATGATGAAGAAGAGCTTTATTTGGCCCAAGAGAGCTTTCA 420
|
|
|
Db 361 GACATCTATGCTATGAAGATGATGAAGAAGAGCTTTATTTGGCCCAAGAGAGCTTTCA 420
|
|
|
QY 421 TTTTGTGAGAAAGCGGACATATTTATCTGAAAGCAAGAGCCGCTGATCCCAATTA 480
|
|
|
Db 421 TTTTGTGAGAAAGCGGACATATTTATCTGAAAGCAAGAGCCGCTGATCCCAATTA 480
|
|
|
QY 481 CAGATAGCTTTTCAGGACAAAATCACTTTATCTGTCATGAAATATCACTGAGAGG 540
|
|
|
Db 481 CAGATAGCTTTTCAGGACAAAATCACTTTATCTGTCATGAAATATCACTGAGAGG 540
|
|
|
QY 541 GACTTGCCTGATCTTTGAATATGATGAGGACCAATTAAGTAAACCTGATACAGTTT 600
|
|
|
Db 541 GACTTGCCTGATCTTTGAATATGATGAGGACCAATTAAGTAAACCTGATACAGTTT 600
|
|
|
QY 601 TACCTAGCTAGCTGATTTTGGCTGTTTCAACGCTTCACTGATGAGTACGTCATCGA 660
|
|
|
Db 601 TACCTAGCTAGCTGATTTTGGCTGTTTCAACGCTTCACTGATGAGTACGTCATCGA 660
|
|
|
QY 661 GACATCAAGCTGAGAAATTTCTGTTGACCGACAGAGACATCAAGCTGAGATTTT 720
|
|
|
Db 661 GACATCAAGCTGAGAAATTTCTGTTGACCGACAGAGACATCAAGCTGAGATTTT 720
|
|
|
QY 721 GGATCTGCGGAAAAATGAATTTCAACAAGATGTTGAATGCAAACTCCGATGGAGC 780
|
|
|
Db 721 GGATCTGCGGAAAAATGAATTTCAACAAGATGTTGAATGCAAACTCCGATGGAGC 780
|
|
|
QY 781 CCAGATTACATGAGCTCCCTGAAGTCTGATCTGTGATGAACGGAGATGAAAAAGCACTAC 840
|
|
|
Db 781 CCAGATTACATGAGCTCCCTGAAGTCTGATCTGTGATGAACGGAGATGAAAAAGCACTAC 840
|
|
|
QY 841 GGCCTGGACTGTGATCTGTGATGATGAGGCGTGAATTCCTATGATGATTTATGAGAGA 900
|
|
|
Db 841 GGCCTGGACTGTGATCTGTGATGATGAGGCGTGAATTCCTATGATGATTTATGAGAGA 900
|
|
|
QY 901 TCCCTCTGAGAGGAGACCTGCGCAAGACCTTCAATTAAGATTAAGATTTTCCAGCG 960
|
|
|
Db 901 TCCCTCTGAGAGGAGACCTGCGCAAGACCTTCAATTAAGATTAAGATTTTCCAGCG 960
|
|
|
QY 961 TTTTGTGAATTTTCCAGATGACCCCAAGTGAAGTACTTTCTGATCTGATTTCAAAGC 1020
|
|
|
Db 961 TTTTGTGAATTTTCCAGATGACCCCAAGTGAAGTACTTTCTGATCTGATTTCAAAGC 1020
|
|
|
QY 1021 TTTGTTGCGGCAAGAAAGAGATGTAAGTGTGAAGTCTTTGCTGCCATCCTTTCTTC 1080
|
|
|
Db 1021 TTTGTTGCGGCAAGAAAGAGATGTAAGTGTGAAGTCTTTGCTGCCATCCTTTCTTC 1080
|
|
|
QY 1081 TCTAAATTTGACTGGAACAACTTCGTAACTCTCTCCCTCCCTGCTTCCCACTCAAG 1140
|
|
|
Db 1081 TCTAAATTTGACTGGAACAACTTCGTAACTCTCTCCCTCCCTGCTTCCCACTCAAG 1140
|
|
|
QY 1141 TCTGAGATGACACTCAATTTTGAATGAACAGAGAAAGATTCGTGGGTTTCACTCTC 1200
|
|
|
Db 1141 TCTGAGATGACACTCAATTTTGAATGAACAGAGAAAGATTCGTGGGTTTCACTCTC 1200
|
|
|
QY 1201 CCGTGCAGCTGAAGCCCTCAAGCTTCTCGGATGAAGAACTGCCTGTTGTGGGTTTTCG 1260
|
|
|
Db 1201 CCGTGCAGCTGAAGCCCTCAAGCTTCTCGGATGAAGAACTGCCTGTTGTGGGTTTTCG 1260
|
|
|
QY 1261 TTAAGAAAGCACTGGGGATCTTGGTGAATCTGAATCTGTTGTGTGGGTTCTGACTCC 1320
|
|
|
Db 1261 TTAAGAAAGCACTGGGGATCTTGGTGAATCTGAATCTGTTGTGTGGGTTCTGACTCC 1320
|
|
|
QY 1321 CCTGCAAGACTAGCTCCATGGAAGAAATTTCTCATAAAGCAAGAGCTTCAAGAC 1380
|
|
|
Db 1321 CCTGCAAGACTAGCTCCATGGAAGAAATTTCTCATAAAGCAAGAGCTTCAAGAC 1380
|
|
|
QY 1381 TCTCAGGACAAGTGTCAAGAATGAGACGAAGATGACCCGTTTACATCGAGAGTGTCA 1440
|
|
|
Db 1381 TCTCAGGACAAGTGTCAAGAATGAGACGAAGATGACCCGTTTACATCGAGAGTGTCA 1440
|
|
|
QY 1441 GAGGTGAGGCTGTGCTTATGTCAGAAAGAGGTGAGCTGAAGGCTCTGAGACTCAGAGA 1500
|
|
|

Db 1441 GAGGTGAGGCTGTGCTTATGTCAGAAAGAGTGAAGTGAAGGCTCTGAGACTCAGAGA 1500
|
|
|
QY 1501 TCCCTCTGAGAGCAGACCTTGTCTAACCAATCAAGATGATGATAGCTTTAAAGGAAAGT 1560
|
|
|
Db 1501 TCCCTCTGAGAGCAGACCTTGTCTAACCAATCAAGATGATGATAGCTTTAAAGGAAAGT 1560
|
|
|
QY 1561 TTTGAGCAAGCACCGATGAGAGTGTCCAGAGAGATGACAAAGCACTGACGCTTCTCAT 1620
|
|
|
Db 1561 TTTGAGCAAGCACCGATGAGAGTGTCCAGAGAGATGACAAAGCACTGACGCTTCTCAT 1620
|
|
|
QY 1621 GATATCAGAGAGACGAGCTGAGTGAATCACTGTAAGAGGATCTTGTCTGAGCAAGAGA 1740
|
|
|
Db 1621 GATATCAGAGAGACGAGCTGAGTGAATCACTGTAAGAGGATCTTGTCTGAGCAAGAGA 1740
|
|
|
QY 1681 GTGGAAGAAATGAGGTTGATGATGAATCACTGTAAGAGGATCTTGTCTGAGCAAGAGA 1740
|
|
|
Db 1681 GTGGAAGAAATGAGGTTGATGATGAATCACTGTAAGAGGATCTTGTCTGAGCAAGAGA 1740
|
|
|
QY 1741 CCGAGTATCTTACGATCTGAGCTGAGAGAGTCTCGGCTTGTCTGTAAGATTTCAAG 1800
|
|
|
Db 1741 CCGAGTATCTTACGATCTGAGCTGAGAGAGTCTCGGCTTGTCTGTAAGATTTCAAG 1800
|
|
|
QY 1801 CCGAAGCCGACAGATGTCAAGCTTAACTGTTGAAGCTTAAGATCAAGGAAAGCTGAA 1860
|
|
|
Db 1801 CCGAAGCCGACAGATGTCAAGCTTAACTGTTGAAGCTTAAGATCAAGGAAAGCTGAA 1860
|
|
|
QY 1861 GTGGAAGAAATATGAGAACTGAGAAAGATCAATGCTGAGAGAGGCTCAAAATTCAGAG 1920
|
|
|
Db 1861 GTGGAAGAAATATGAGAACTGAGAAAGATCAATGCTGAGAGAGGCTCAAAATTCAGAG 1920
|
|
|
QY 1921 CTCCAAGAGAACTGAGAGAGCTGTAAAGCCAGACCGAGGCCACCGAGCTGTGAG 1980
|
|
|
Db 1921 CTCCAAGAGAACTGAGAGAGCTGTAAAGCCAGACCGAGGCCACCGAGCTGTGAG 1980
|
|
|
QY 1981 AATATCCGACAGCAAAAGAGCGAGCCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2040
|
|
|
Db 1981 AATATCCGACAGCAAAAGAGCGAGCCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2040
|
|
|
QY 2041 GATTTCTTGAAGGCAATCAAGAAAGAGCTGTGGAAGCTGAGGAGAGCGGCCATTCCTG 2100
|
|
|
Db 2041 GATTTCTTGAAGGCAATCAAGAAAGAGCTGTGGAAGCTGAGGAGAGCGGCCATTCCTG 2100
|
|
|
QY 2101 GAGAACAGAGTAAAGAGCTAAGACATGAGAGCTGAGAGAAACAGACTGAAGATGAC 2160
|
|
|
Db 2101 GAGAACAGAGTAAAGAGCTAAGACATGAGAGCTGAGAGAAACAGACTGAAGATGAC 2160
|
|
|
QY 2161 ATTCAGACAAATCCCAAGATCCAGAGATGAGCTGATTAATTTTGAAGCTCGAAGAG 2220
|
|
|
Db 2161 ATTCAGACAAATCCCAAGATCCAGAGATGAGCTGATTAATTTTGAAGCTCGAAGAG 2220
|
|
|
QY 2221 AAAATCCGAGAGGCCCAAGCTCAAGCCAGACCTAAGAGTGAACCTGAAGAGAGAG 2280
|
|
|
Db 2221 AAAATCCGAGAGGCCCAAGCTCAAGCCAGACCTAAGAGTGAACCTGAAGAGAGAG 2280
|
|
|
QY 2281 CAGCACTATGAAGAAAGATTAAGTGTGGAACAATCAAGATTAAGAAAGACCTGCTGAC 2340
|
|
|
Db 2281 CAGCACTATGAAGAAAGATTAAGTGTGGAACAATCAAGATTAAGAAAGACCTGCTGAC 2340
|
|
|
QY 2341 AAGAGACACTGAGAGACATGATGACAGACACGAGAGAGAGGCCCATGAGAGAGGCAAA 2400
|
|
|
Db 2341 AAGAGACACTGAGAGACATGATGACAGACACGAGAGAGAGGCCCATGAGAGAGGCAAA 2400
|
|
|
QY 2401 AATTCAGAGAAACAGAGCCGATGATCAATGCTTAAGATTTTCAAGATCAAGATCCCTGAA 2460
|
|
|
Db 2401 AATTCAGAGAAACAGAGCCGATGATCAATGCTTAAGATTTTCAAGATCAAGATCCCTGAA 2460
|
|
|
QY 2461 CAGAGATTTGTGAACTGTCTGAAGCCAAATTAAGTGAAGCAATAGCAGTCTTTTACC 2520
|
|
|
Db 2461 CAGAGATTTGTGAACTGTCTGAAGCCAAATTAAGTGAAGCAATAGCAGTCTTTTACC 2520
|
|
|
QY 2521 CAAAGAAATGAGAGGCCCAAGAGAGATGATTTCTGAATCTGAGCAACAGAAATTTTAC 2580
|
|
|

Dh 2521 CAAAGACATGAAAGCCCAAGAAAGATGATTTCTGAACTCAGGCAACGAAATTTTAC 2580
Qy 2581 CTGGAGACAGAGGCTGGGAAGTTGGAGGCCAGAACCGAAAACTGGAGAGCAGCTGGAG 2640
Dh 2581 CTGGAGACAGAGGCTGGGAAGTTGGAGGCCAGAACCGAAAACTGGAGAGCAGCTGGAG 2640
Qy 2641 AAGATCAGCCACCAAGACCAAGTGAACAAGATCGGCTGCTGGAATCTGGAGACAAAGTTG 2700
Dh 2641 AAGATCAGCCACCAAGACCAAGTGAACAAGATCGGCTGCTGGAATCTGGAGACAAAGTTG 2700
Qy 2701 CGGAGGTCAGTCTGAGAGCAAGAGACAGAACTGAGAGCTCAAGCCGACAGCTCAAGAG 2760
Dh 2701 CGGAGGTCAGTCTGAGAGCAAGAGACAGAACTGAGAGCTCAAGCCGACAGCTCAAGAG 2760
Qy 2761 CTACAGCTCTCCCTGAGAGCGGAGTCAAGTGTGACAGCCCTGACGAGCTGACAGGCG 2820
Dh 2761 CTACAGCTCTCCCTGAGAGCGGAGTCAAGTGTGACAGCCCTGACGAGCTGACAGGCG 2820
Qy 2821 GCCCTGAGAGCCAGCTTTCGCAAGCGAAGACAGAGCTGAGAGACCAAGAGAGCT 2880
Dh 2821 GCCCTGAGAGCCAGCTTTCGCAAGCGAAGACAGAGCTGAGAGACCAAGAGAGCT 2880
Qy 2881 GAAGAGAGATCCAGGACCTCAAGGCACTAGAGATGAAATCCAGCGGAAATTTGAGCT 2940
Dh 2881 GAAGAGAGATCCAGGACCTCAAGGCACTAGAGATGAAATCCAGCGGAAATTTGAGCT 2940
Qy 2941 CTTGCTAACAGCTGTAATCTGTAATCAAGAGCTGAGAGAGCAGCTAAACAGCTGACGAG 3000
Dh 2941 CTTGCTAACAGCTGTAATCTGTAATCAAGAGCTGAGAGAGCAGCTAAACAGCTGACGAG 3000
Qy 3001 GACAAAGCTGAATCTCAACAAACAACTTCTGTCGAAACAACTGATGAGGCTTCT 3060
Dh 3001 GACAAAGCTGAATCTCAACAAACAACTTCTGTCGAAACAACTGATGAGGCTTCT 3060
Qy 3061 GGGCCCAACAGAGATTTGTAACAATGGAGAGTGAAGTGAACATCTCCCGGGAGATC 3120
Dh 3061 GGGCCCAACAGAGATTTGTAACAATGGAGAGTGAAGTGAACATCTCCCGGGAGATC 3120
Qy 3121 ACGGAACGAGAGATGAGCTTACAGCAGAGAGCAAGCAAGTGAAGGCTCTGAAGACACG 3180
Dh 3121 ACGGAACGAGAGATGAGCTTACAGCAGAGAGCAAGCAAGTGAAGGCTCTGAAGACACG 3180
Qy 3181 TGCAACATGCTGAGAGAAAGATGATTTGGAGGCTTAAACGATGAGCTGTAGAA 3240
Dh 3181 TGCAACATGCTGAGAGAAAGATGATTTGGAGGCTTAAACGATGAGCTGTAGAA 3240
Qy 3241 AAAGAGCGGAGTGGAGAGGCTTGAAGAGCGTCTGGGTGATGAGAAATCCAGTTGAG 3300
Dh 3241 AAAGAGCGGAGTGGAGAGGCTTGAAGAGCGTCTGGGTGATGAGAAATCCAGTTGAG 3300
Qy 3301 TGTGGGTTTGGAGAGCTGCAAGAGATGCTGACACCGAGAAACAGAGCGGCGAGAGC 3360
Dh 3301 TGTGGGTTTGGAGAGCTGCAAGAGATGCTGACACCGAGAAACAGAGCGGCGAGAGC 3360
Qy 3361 GATCAGCGGATCAACGAGTCTGCGCAGGTGGTGGAGCTGAGTGAAGAGACAAAGGCT 3420
Dh 3361 GATCAGCGGATCAACGAGTCTGCGCAGGTGGTGGAGCTGAGTGAAGAGACAAAGGCT 3420
Qy 3421 GAGATTCTCGCTGAGAGAGGCTTCAAGAGCGAAGCTGAAGGCGGAGAGCTCTCT 3480
Dh 3421 GAGATTCTCGCTGAGAGAGGCTTCAAGAGCGAAGCTGAAGGCGGAGAGCTCTCT 3480
Qy 3481 GACAAAGCTCAATGAGCTTGAAGAGAGATGATGATTTGAATGAATGCCGAGAGCTTA 3540
Dh 3481 GACAAAGCTCAATGAGCTTGAAGAGAGATGATGATTTGAATGAATGCCGAGAGCTTA 3540
Qy 3541 CAGCAGAGCTGAGAGCTGAAGAGAGTCAAAACGAGAGCTTCTGGAAGAGAAACCAA 3600
Dh 3541 CAGCAGAGCTGAGAGCTGAAGAGAGTCAAAACGAGAGCTTCTGGAAGAGAAACCAA 3600
Qy 3601 TTACAGCAGAGAGTGAAGCTGCAAGAAATCAATCTTTCGCTGATCAAGAGCTGCA 3660
Dh 3601 TTACAGCAGAGAGTGAAGCTGCAAGAAATCAATCTTTCGCTGATCAAGAGCTGCA 3660

Qy 3661 GAAGCTTGAATGGGCTGATCTTACTGAGACAGAAAGAGTGAATCTGAGATACAGCTG 3720
Dh 3661 GAAGCTTGAATGGGCTGATCTTACTGAGACAGAAAGAGTGAATCTGAGATACAGCTG 3720
Qy 3721 GAAACATTCAGGTTCTTATTCCTATGAAAGGTTGAAGTGAAGGACATTTCTCA 3780
Dh 3721 GAAACATTCAGGTTCTTATTCCTATGAAAGGTTGAAGTGAAGGACATTTCTCA 3780
Qy 3781 CAAACCAATCTATTTGATTTTCTGCAAGCCAAATGAGACCACTGCTAAAAAGAAAG 3840
Dh 3781 CAAACCAATCTATTTGATTTTCTGCAAGCCAAATGAGACCACTGCTAAAAAGAAAG 3840
Qy 3841 GTTCTCTGAGTACAAATGAGCTGAGCTGAGCTGAGAGAGAGAAAGTCTGCTGCA 3900
Dh 3841 GTTCTCTGAGTACAAATGAGCTGAGCTGAGCTGAGAGAGAGAAAGTCTGCTGCA 3900
Qy 3901 GAGCTTGAAGAGGCTTCTGAGAGACCCGCAATGAGAGCTGAGAGCTGAGAGGAGAGCT 3960
Dh 3901 GAGCTTGAAGAGGCTTCTGAGAGACCCGCAATGAGAGCTGAGAGCTGAGAGGAGAGCT 3960
Qy 3961 GCCCAGCGCAAGAGAGAGAGACCAACCACTCAACGAGCAGCAGCAGCAGCAGCAG 4020
Dh 3961 GCCCAGCGCAAGAGAGAGAGACCAACCACTCAACGAGCAGCAGCAGCAGCAGCAG 4020
Qy 4021 ATGCGCATGTCGCGCATGTCGCGTGCAGAGACCAAGCCAGTCCATGAGCTGCTG 4080
Dh 4021 ATGCGCATGTCGCGCATGTCGCGTGCAGAGACCAAGCCAGTCCATGAGCTGCTG 4080
Qy 4081 GCCCGCGATTCAGCGCAGAAAGAGATCTTCACTCAAGAGAAATTTAGTCGCGCTT 4140
Dh 4081 GCCCGCGATTCAGCGCAGAAAGAGATCTTCACTCAAGAGAAATTTAGTCGCGCTT 4140
Qy 4141 AAGGAAGCATGACCAACATATCTCAACGATTCAGAGTGAAGTGAAGTGAAGC 4200
Dh 4141 AAGGAAGCATGACCAACATATCTCAACGATTCAGAGTGAAGTGAAGTGAAGC 4200
Qy 4201 ACAAGTGTGCTGTGTCTTGGATACGCTGACCTTTGAGAGCGCAGCATCCAATGTCTC 4260
Dh 4201 ACAAGTGTGCTGTGTCTTGGATACGCTGACCTTTGAGAGCGCAGCATCCAATGTCTC 4260
Qy 4261 GAATGTCAAGTATGTGTCAACCCCAAGTCTTCAACGCTTCCAGCCACTGGGGCTTG 4320
Dh 4261 GAATGTCAAGTATGTGTCAACCCCAAGTCTTCAACGCTTCCAGCCACTGGGGCTTG 4320
Qy 4321 CTTGCTGAATATGCAACACTTCAACGAGGCTTCTGCGGTGACAAATGAATCACTCCCA 4380
Dh 4321 CTTGCTGAATATGCAACACTTCAACGAGGCTTCTGCGGTGACAAATGAATCACTCCCA 4380
Qy 4381 GGTCTCAGACCAAGAGGCTCAGAGAGCTTTCGACCTGGAAGGTTGATGAAGTGC 4440
Dh 4381 GGTCTCAGACCAAGAGGCTCAGAGAGCTTTCGACCTGGAAGGTTGATGAAGTGC 4440
Qy 4441 AGGAATTAACAAACGAGAGAGGCTGAGAGAGTCAATTTCTCTGAGAGATCA 4500
Dh 4441 AGGAATTAACAAACGAGAGAGGCTGAGAGAGTCAATTTCTCTGAGAGATCA 4500
Qy 4501 AAAGTCTCAATTTATGCAATGAAGCAGAGAGTGAAGAGGCGGAGAGAGATTT 4560
Dh 4501 AAAGTCTCAATTTATGCAATGAAGCAGAGAGTGAAGAGGCGGAGAGAGATTT 4560
Qy 4561 GAGCTGTGCTTCCCGAGCGGAGATGATCTTTCGAGTGTGCTGCTTCCGAACTC 4620
Dh 4561 GAGCTGTGCTTCCCGAGCGGAGATGATCTTTCGAGTGTGCTGCTTCCGAACTC 4620
Qy 4621 GCAATATACGCAAGAGAGAGTCCATACATCTGAAGATGAGATCTCACCCGACAC 4680
Dh 4621 GCAATATACGCAAGAGAGAGTCCATACATCTGAAGATGAGATCTCACCCGACAC 4680
Qy 4681 AACTGTGAGCGCGAGAGAGCTTACTTGTGAGTCCAGCTTCCCTGACAAACAGGCG 4740
Dh 4681 AACTGTGAGCGCGAGAGAGCTTACTTGTGAGTCCAGCTTCCCTGACAAACAGGCG 4740

4741 TGGGTCAACCGCTTGAATCAGTTGTGGCAGGTGGGAGATTCTTAGGGAAAAACGACAA 4800
4741 TGGGTCAACCGCTTGAATCAGTTGTGGCAGGTGGGAGATTCTTAGGGAAAAACGACAA 4800
4801 GCTGATGTAACTGCTTGAAGACTCCCTGCTGAATCTGAAGGTGATGACCTGTAGAC 4860
4801 GCTGATGTAACTGCTTGAAGACTCCCTGCTGAATCTGAAGGTGATGACCTGTAGAC 4860
4861 ATGAATGCAAGCTGCTTGAAGACTCCCTGCTGAATCTGAAGGTGATGACCTGTAGAC 4920
4861 ATGAATGCAAGCTGCTTGAAGACTCCCTGCTGAATCTGAAGGTGATGACCTGTAGAC 4920
4921 TACGCCCTGAATGCTTGAAGAACTCCCTAAGCCATGCTCCAGGAATTGAGAGCTTTC 4980
4921 TACGCCCTGAATGCTTGAAGAACTCCCTAAGCCATGCTCCAGGAATTGAGAGCTTTC 4980
4981 CAATTTATATATTAACAAGACTCTGAGAACTGATGATGAGAGAGAGAGAGAGAGAG 5040
4981 CAATTTATATTAACAAGACTCTGAGAACTGATGATGAGAGAGAGAGAGAGAGAGAG 5040
5041 CTGTGCTTGTGAGAGTGAAGAAAGTGAACAGTCCCTGGCCAGTCCACCTGCTGCC 5100
5041 CTGTGCTTGTGAGAGTGAAGAAAGTGAACAGTCCCTGGCCAGTCCACCTGCTGCC 5100
5101 CAGCCCGACATCTCAACCAATTTTGAAGCTGTCAAGGCTGCCACTTGTGGGGCA 5160
5101 CAGCCCGACATCTCAACCAATTTTGAAGCTGTCAAGGCTGCCACTTGTGGGGCA 5160
5161 GGCAGATTGAGAAAGGCTCTGATCTGTGAGAGCTGAGCAAGTCCAGCAAGTCTGCTC 5220
5161 GGCAGATTGAGAAAGGCTCTGATCTGTGAGAGCTGAGCAAGTCCAGCAAGTCTGCTC 5220
5221 CGCTAACCAAAAACCTCAGCAAAATGTCATCGGAAAGATGAGAGCTCAGAGCCC 5280
5221 CGCTAACCAAAAACCTCAGCAAAATGTCATCGGAAAGATGAGAGCTCAGAGCCC 5280
5281 TGCAGCTGATCACTTCAACCAATTCAGATCTCAATGGAACCAATAATCTAGAA 5340
5281 TGCAGCTGATCACTTCAACCAATTCAGATCTCAATGGAACCAATAATCTAGAA 5340
5341 ATGACATGAGAGAGTACAGCTCGAGAGATCTGAGATTAAGATGACATTCCTGGCA 5400
5341 ATGACATGAGAGAGTACAGCTCGAGAGATCTGAGATTAAGATGACATTCCTGGCA 5400
5401 CCGTGTGTGTTGGCCCTCTTCAACAGCTTCTGTCTCAATCTGTGAGTGAACAGC 5460
5401 CCGTGTGTGTTGGCCCTCTTCAACAGCTTCTGTCTCAATCTGTGAGTGAACAGC 5460
5461 GCGAGGACAGAGAGAGTACCTGCTGTGTTTCAAGATTTGAGATGCTGAGATTCT 5520
5461 GCGAGGACAGAGAGAGTACCTGCTGTGTTTCAAGATTTGAGATGCTGAGATTCT 5520
5521 TACGAGAGAGTACGAGAGAGTACGATCTCAATGAGTGCCTTACCTTGGCTTGGC 5580
5521 TACGAGAGAGTACGAGAGAGTACGATCTCAATGAGTGCCTTACCTTGGCTTGGC 5580
5581 TACGAGAGAGTACGATCTGTTTGTGACCACTTCACTCACTGAGATTAATTGAGTCA 5640
5581 TACGAGAGAGTACGATCTGTTTGTGACCACTTCACTCACTGAGATTAATTGAGTCA 5640
5641 GAGAGCTCTCAGAGAGAGAGTCCCTGCGAGAGTACTGAGATCCCGAGAGAGAGGCTC 5700
5641 GAGAGCTCTCAGAGAGAGAGTCCCTGCGAGAGTACTGAGATCCCGAGAGAGAGGCTC 5700
5701 CTGGAGCCCTGCAATTTCTCAGAGAGATTTACTTGGAGTCTTCAACAGAGATTAATTA 5760
5701 CTGGAGCCCTGCAATTTCTCAGAGAGATTTACTTGGAGTCTTCAACAGAGATTAATTA 5760
5761 AGGATCATTTGCTGCAAGGAGAACTCTGTAAGAGTCCGAGTCACTGAGACACCGGGC 5820
5761 AGGATCATTTGCTGCAAGGAGAACTCTGTAAGAGTCCGAGTCACTGAGACACCGGGC 5820
5821 CCGTCAACTCTCCGAGAGAGCCCAACAGAGAGAGCCCACTGACGATCAACAGAGACATC 5880

5821 CCGTCAACTCTCCGAGAGAGCCCAACAGAGAGAGCCCACTGACGATCAACAGACACATC 5880
5881 ACNAGAGGCTGAGCTCTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5940
5881 ACNAGAGGCTGAGCTCTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5940
5941 AGCAGACCCAGCTTACCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6000
5941 AGCAGACCCAGCTTACCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6000
6001 CCGCCCTGAG 6060
6001 CCGCCCTGAG 6060
6061 CCGCCAGAGCTGTTTGAAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6120
6061 CCGCCAGAGCTGTTTGAAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6120
6121 CCGCTGTCCAGGTGAACAAGATCTGGACAGATCTTCAATATTA 6165
6121 CCGCTGTCCAGGTGAACAAGATCTGGACAGATCTTCAATATTA 6165

RESULT 2
US-10-028-946-3
; Sequence 3, Application US/10028946
; Patent No. 6734009
; GENERAL INFORMATION:
; APPLICANT: Yu, Xuanchuan
; APPLICANT: Miranda, Maricar
; APPLICANT: Fiddie, Carl Johan
; TITLE OF INVENTION: No. 6734009el Human Kinases and Polynucleotides Encoding the Same
; FILE REFERENCE: LEX-0289-USA
; CURRENT APPLICATION NUMBER: US/10/028, 946
; PRIOR FILING DATE: 2001-12-20
; PRIOR APPLICATION NUMBER: US 60/258, 335
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 3
; LENGTH: 5877
; TYPE: DNA
; ORGANISM: homo sapiens
; US-10-028-946-3

Query Match 94.7%; Score 5837; DB 4; Length 5877;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 5837; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGTTGAAGTTCAAAATATGAGAGAGAGATCTTGAATGCTGAGTGTGAGACCAT 60
DB 1 ATGTTGAAGTTCAAAATATGAGAGAGAGATCTTGAATGCTGAGTGTGAGACCAT 60
QY 61 GCCAGCCGGGCTCCAGGCTGAATCTTTCAGAGGAGAAACCACTTTATGACTCA 120
DB 61 GCCAGCCGGGCTCCAGGCTGAATCTTTCAGAGGAGAAACCACTTTATGACTCA 120
QY 121 CAGCAGATGTCTCTCTTTCAGAGAGAGATTAAGATGCTCTTGTGTTCTTTGAA 180
DB 121 CAGCAGATGTCTCTCTTTCAGAGAGAGATTAAGATGCTCTTGTGTTCTTTGAA 180
QY 181 GAATCAGTCAAGCTGCTGTATGAAATTAAGACGTGAGCACTTTGTCCGGAAGAT 240
DB 181 GAATCAGTCAAGCTGCTGTATGAAATTAAGACGTGAGCACTTTGTCCGGAAGAT 240
QY 241 TCCGACACCATAGCTGAGTTACAGAGCTCCAGCTTCCGGAAGAGACTTGAAGTCA 300
DB 241 TCCGACACCATAGCTGAGTTACAGAGCTCCAGCTTCCGGAAGAGACTTGAAGTCA 300
QY 301 AGTCTTGAAGTGTGTTGTTCACTTGTGTAAGTCAAGTGTGTAAGAGAGAAACCGGG 360
DB 301 AGTCTTGAAGTGTGTTGTTCACTTGTGTAAGTCAAGTGTGTAAGAGAGAAACCGGG 360

QY	1441	GAGGTGAAGGCTGTGCTTACTCAGAAAGAGGTGAGCTGAAGGCTCTGAGACTCAGAGA	1500
Db	1441	GAGGTGAAGGCTGTGCTTACTCAGAAAGAGGTGAGCTGAAGGCTCTGAGACTCAGAGA	1500
QY	1501	TTCCCTCTGAGAGCAGGACCTTGCTACCTCATCAGAAATGCAGTATGCTTAAAGCGAAGT	1560
Db	1501	TTCCCTCTGAGAGCAGGACCTTGCTACCTCATCAGAAATGCAGTATGCTTAAAGCGAAGT	1560
QY	1561	TTGGAGCAGACGAGATGAGAGGTGTCCCAGAGAGATGACAAAGCACTGCACTTCTCAAT	1620
Db	1561	TTGGAGCAGACGAGATGAGAGGTGTCCCAGAGAGATGACAAAGCACTGCACTTCTCAAT	1620
QY	1621	GATATCAAGAGCAGAGCCGGAAGCTCCAGAAATCAAAAGCAGAGATGCCAGGCTCA	1680
Db	1621	GATATCAAGAGCAGAGCCGGAAGCTCCAGAAATCAAAAGCAGAGATGCCAGGCTCA	1680
QY	1681	GTGGAAGAAATGAGGTTGATGATGAATCAAGTTGGAAAGAGATCTTGTCTCAGCAAGAG	1740
Db	1681	GTGGAAGAAATGAGGTTGATGATGAATCAAGTTGGAAAGAGATCTTGTCTCAGCAAGAG	1740
QY	1741	CGAGATGATCTCTACGAATCTGAGCTGAGAGATCTCGGCTTGCTGTAAGAAATTCAG	1800
Db	1741	CGAGATGATCTCTACGAATCTGAGCTGAGAGATCTCGGCTTGCTGTAAGAAATTCAG	1800
QY	1801	CGAAAGCGACAGATGTCAAGATAAACTGTTAAGGCTTAAGATCAAGGAAAGCTGAA	1860
Db	1801	CGAAAGCGACAGATGTCAAGATAAACTGTTAAGGCTTAAGATCAAGGAAAGCTGAA	1860
QY	1861	GTGGAGAAATTTGTGGAAACTGAGAAAGATTCATATCTGACAGCAGCTCAAAATTCAGAG	1920
Db	1861	GTGGAGAAATTTGTGGAAACTGAGAAAGATTCATATCTGACAGCAGCTCAAAATTCAGAG	1920
QY	1921	CTCCAGAGAAACTCGAGAAAGGCTGTAAAGCCAGCAGGAGCCACCGAGCTGCTGAG	1980
Db	1921	CTCCAGAGAAACTCGAGAAAGGCTGTAAAGCCAGCAGGAGCCACCGAGCTGCTGAG	1980
QY	1981	AATATCCGCCAGGCAAGAGCGAGCCGAGAGGAGCTGAGAAAGCTGCAGAACCGAGAG	2040
Db	1981	AATATCCGCCAGGCAAGAGCGAGCCGAGAGGAGCTGAGAAAGCTGCAGAACCGAGAG	2040
QY	2041	GATTCTTCTGAGGCACTCAGAAAGAGCTGTGTGAAGCTGAGGAAAGCGCCGCAATTTCTG	2100
Db	2041	GATTCTTCTGAGGCACTCAGAAAGAGCTGTGTGAAGCTGAGGAAAGCGCCGCAATTTCTG	2100
QY	2101	GAGAAACAAGGTAAAGAGACCTAGAGACCATGAGGCTGTAGAAACAGACTGAGAGATGAC	2160
Db	2101	GAGAAACAAGGTAAAGAGACCTAGAGACCATGAGGCTGTAGAAACAGACTGAGAGATGAC	2160
QY	2161	ATCCAGACAATAATCCCAACAGATCCAGCAGATGGCTGTATAAATTTCTGAGCTCGAAGAG	2220
Db	2161	ATCCAGACAATAATCCCAACAGATCCAGCAGATGGCTGTATAAATTTCTGAGCTCGAAGAG	2220
QY	2221	AAACATGGGAGGCCCAAGTCTCAGGCCACCACTGAAATGCACTGTAACCTGTAACGAAAGAG	2280
Db	2221	AAACATGGGAGGCCCAAGTCTCAGGCCACCACTGAAATGCACTGTAACCTGTAACGAAAGAG	2280
QY	2281	CAGACATATGAGGAAAGATTTAAAGTGTGGAACAATCAGATTAAAGAAAGACTGAGCTGAC	2340
Db	2281	CAGACATATGAGGAAAGATTTAAAGTGTGGAACAATCAGATTAAAGAAAGACTGAGCTGAC	2340
QY	2341	AAGAGACACACTGAGAAACATGATGACAGAGACGAGAGAGAGGCCCATGAGAGGGCCAA	2400
Db	2341	AAGAGACACACTGAGAAACATGATGACAGAGACGAGAGAGAGGCCCATGAGAGGGCCAA	2400
QY	2401	ATTCTCAGCCGAACGAAAGCGATGATCAATGCTATGGAATTCCAAATCAGATCCCTGGA	2460
Db	2401	ATTCTCAGCCGAACGAAAGCGATGATCAATGCTATGGAATTCCAAATCAGATCCCTGGA	2460
QY	2461	CAGAGATTTGGGAACGTCTGGAAGCCCAATTAACCTTGACCAAAATGACACTCTTTTACC	2520
Db	2461	CAGAGATTTGGGAACGTCTGGAAGCCCAATTAACCTTGACCAAAATGACACTCTTTTACC	2520
QY	2521	CAAGGAACATGAAGGCCCAAGAGAGANTATTTCTGMACTCAGGCAACGAAATTTTAC	2580

Db 2521 CAAAGGAACATGAGAGGCCCAAGAGAGATGATTTTCTGAACTCAGGCCAACAAGAAATTTTAC 2580
Qy 2581 CTGGAACACAGGCTGGAGATTGAGGCCCAAGAACCGAAATCTGGAGAGCAGCTGAG 2640
Db 2581 CTGGAAGACACAGGCTGGAGATTGAGGCCCAAGAACCGAAATCTGGAGAGCAGCTGAG 2640
Qy 2641 AAGATCAGCCCAACCAAGACCAACAGTGAACAAGATCGGCTGCTGGAACCTGGAGACAAGATTG 2700
Db 2641 AAGATCAGCCCAACCAAGACCAACAGTGAACAAGATCGGCTGCTGGAACCTGGAGACAAGATTG 2700
Qy 2701 CGGAGGCTCAGTCTAGAGCAACGAGAGCAAGAACTGAGCTCAGACGCGCAGCTCAGAG 2760
Db 2701 CGGAGGCTCAGTCTAGAGCAACGAGAGCAAGAACTGAGAGCTCAGACGCGCAGCTCAGAG 2760
Qy 2761 CTACAGCTCTCCCTGCAAGAGCGGAGTCAAGTTGACAGCCCTGACAGCTGCAAGGCG 2820
Db 2761 CTACAGCTCTCCCTGCAAGAGCGGAGTCAAGTTGACAGCCCTGACAGCTGCAAGGCG 2820
Qy 2821 GCCCTGAGAGCCAGCTTCCGCAAGCGCAAGACAGAGCTGGAAGAGACACAGCAGAACT 2880
Db 2821 GCCCTGAGAGCCAGCTTCCGCAAGCGCAAGACAGAGCTGGAAGAGACACAGCAGAACT 2880
Qy 2881 GAAGAGAGATCCAGGCACTCAGGCACTAGAGATGAATCCAGGCGCAATTTGATGCT 2940
Db 2881 GAAGAGAGATCCAGGCACTCAGGCACTAGAGATGAATCCAGGCGCAATTTGATGCT 2940
Qy 2941 CTTCTGAACAGCTGATCTGTATCAAGACCTGGAGAGAGCTAAACAGCTGACCGAG 3000
Db 2941 CTTCTGAACAGCTGATCTGTATCAAGACCTGGAGAGAGCTAAACAGCTGACCGAG 3000
Qy 3001 GACAAGCTGAATCTCAACCAACCAAACTTTCTATTGTTCAAACCACTGATGAGGCTTCT 3060
Db 3001 GACAAGCTGAATCTCAACCAACCAAACTTTCTATTGTTCAAACCACTGATGAGGCTTCT 3060
Qy 3061 GGGCCCAACGACGAGATTGTACAACCTGCAAGTGAAGTGAACCATCTCGCGCGGAGATC 3120
Db 3061 GGGCCCAACGACGAGATTGTACAACCTGCAAGTGAAGTGAACCATCTCGCGCGGAGATC 3120
Qy 3121 ACGGAAACGAGAGATGCACTTACCAAGCAGAGCAACGATGAGGCTCTGAAGACCAAG 3180
Db 3121 ACGGAAACGAGAGATGCACTTACCAAGCAGAGCAACGATGAGGCTCTGAAGACCAAG 3180
Qy 3181 TGACCATGCTGGAAGAAAGAGTCAATGATTTGAGAGCCCTAAACGATGAGCTGTAAG 3240
Db 3181 TGACCATGCTGGAAGAAAGAGTCAATGATTTGAGAGCCCTAAACGATGAGCTGTAAG 3240
Qy 3241 AAAGACGCGACGAGGAGGCTGAGAGAGGCTCTGGGTGATGAGAAATCCAGTTGAG 3300
Db 3241 AAAGACGCGACGAGGAGGCTGAGAGAGGCTCTGGGTGATGAGAAATCCAGTTGAG 3300
Qy 3301 TGTGGGTTTGAAGAGCTGCAAGAGATGCTGACACCGAAGAAACAGACAGGCGGAGAGCC 3360
Db 3301 TGTGGGTTTGAAGAGCTGCAAGAGATGCTGACACCGAAGAAACAGACAGGCGGAGAGCC 3360
Qy 3361 GATCAGCGGATCAACCAAGTCTCCCAAGTGTGAGAGCTGAGCAAGTGAAGCAAGGCT 3420
Db 3361 GATCAGCGGATCAACCAAGTCTCCCAAGTGTGAGAGCTGAGCAAGTGAAGCAAGGCT 3420
Qy 3421 GAGATTCTGCTCTGAGAGAGGCTCAAGAGCAGAACTGAAGGCCAGAGCCCTCTCT 3480
Db 3421 GAGATTCTGCTCTGAGAGAGGCTCAAGAGCAGAACTGAAGGCCAGAGCCCTCTCT 3480
Qy 3481 GACAAGCTCAATGACTGAGAGAAAGAGATGCTATCTTGAATGAAATGAAATCCCGAGCTTA 3540
Db 3481 GACAAGCTCAATGACTGAGAGAAAGAGATGCTATCTTGAATGAAATGAAATCCCGAGCTTA 3540
Qy 3541 CAGCAGAGCTGAGACTGAACAGAGCTCAACAGAGGCTTCTGGAGAGAGCAAGCCAA 3600
Db 3541 CAGCAGAGCTGAGACTGAACAGAGCTCAACAGAGGCTTCTGGAGAGAGCAAGCCAA 3600
Qy 3601 TTACAGCAGCAGATGAGACTGCAAGAAATCAATTTCCGCTGACTCAAGACTGCA 3660
Db 3601 TTACAGCAGCAGATGAGACTGCAAGAAATCAATTTCCGCTGACTCAAGACTGCA 3660

Db 3601 TTACAGCAGCAGATGAGACTGCAAGAAATCAATTTCCGCTGACTCAAGACTGCA 3660
Qy 3661 GAAGCTCTAGATCGGCTGATCTTACTGAGACAGAAAGAGTGAATCTGAGATCAAGCTG 3720
Db 3661 GAAGCTCTAGATCGGCTGATCTTACTGAGACAGAAAGAGTGAATCTGAGATCAAGCTG 3720
Qy 3721 GAAAACTTCAAGTTCTCTATTCTCATGAAAGGTGAAAATGGAAGCACTATTCTCA 3780
Db 3721 GAAAACTTCAAGTTCTCTATTCTCATGAAAGGTGAAAATGGAAGCACTATTCTCA 3780
Qy 3781 CAAACCAATCTATGATTTTCTGCAAGCCCAAAATGGAACCACTGTAAGAAAAAG 3840
Db 3781 CAAACCAATCTATGATTTTCTGCAAGCCCAAAATGGAACCACTGTAAGAAAAAG 3840
Qy 3841 GTTCTCTGAGTACATGAGCTGAAGCTGAGCCCTGAGAGAGAGAAAGCTGCTGTGCA 3900
Db 3841 GTTCTCTGAGTACATGAGCTGAAGCTGAGCCCTGAGAGAGAGAAAGCTGCTGTGCA 3900
Qy 3901 GAGCTGAGAGAGCTTCAAGAAACCCGATGAGCTCCGTCGCGCGGAGAGAACT 3960
Db 3901 GAGCTGAGAGAGCTTCAAGAAACCCGATGAGCTCCGTCGCGCGGAGAGAACT 3960
Qy 3961 GCCCACGGAGAACAGGACCAAGCCCAACCCATCCACGCAAGCCGAGGAGGCAAG 4020
Db 3961 GCCCACGGAGAACAGGACCAAGCCCAACCCATCCACGCAAGCCGAGGAGGCAAG 4020
Qy 4021 ATGCGCATGTCGCACTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 4080
Db 4021 ATGCGCATGTCGCACTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 4080
Qy 4081 GCCCGCCATCCAGCCCGCAAGAAAGAGTCTTCAACTCCAGAGAAATTTAGTCGCTCTT 4140
Db 4081 GCCCGCCATCCAGCCCGCAAGAAAGAGTCTTCAACTCCAGAGAAATTTAGTCGCTCTT 4140
Qy 4141 AAGGAACGATGACCAACCAATTCCTCAACCGATTAAGTGAAGTGAACATGCGAGCC 4200
Db 4141 AAGGAACGATGACCAACCAATTCCTCAACCGATTAAGTGAAGTGAACATGCGAGCC 4200
Qy 4201 ACAAACTGCTGT 4260
Db 4201 ACAAACTGCTGT 4260
Qy 4261 GAATGTCAGTATGATGTGTCAACCCCAAGTCTCAAGTCTTGTGACAGCCAGCTTGTG 4320
Db 4261 GAATGTCAGTATGATGTGTCAACCCCAAGTCTCAAGTCTTGTGACAGCCAGCTTGTG 4320
Qy 4321 CCGTGTGAATATGCCACACACTTCAACCGAGGCTTCTGCGGTGACAAATGGAATCC 4380
Db 4321 CCGTGTGAATATGCCACACACTTCAACCGAGGCTTCTGCGGTGACAAATGGAATCC 4380
Qy 4381 GGTCTTCAGACCAAGAGCCCAAGAGCTTGTGCACTTGTGAGAGGTGATGAGAGTCC 4440
Db 4381 GGTCTTCAGACCAAGAGCCCAAGAGCTTGTGCACTTGTGAGAGGTGATGAGAGTCC 4440
Qy 4441 AAGAAATTAACAAACGAGACAGAGGCTGGAACAGAGAGCTGGAAGAGGATGAAATTT 4500
Db 4441 AAGAAATTAACAAACGAGACAGAGGCTGGAACAGAGAGCTGGAAGAGGATGAAATTT 4500
Qy 4501 AAAGTCTCAATTTAAGACAAATGAACCAAGAGAGCTGGAACAGAGGCTGGAAGATTT 4560
Db 4501 AAAGTCTCAATTTAAGACAAATGAACCAAGAGAGCTGGAACAGAGGCTGGAAGATTT 4560
Qy 4561 GAGCTGTGCTTCCGAGCGGAGATGATCTATCTATCTATCTATCTATCTATCTATCT 4620
Db 4561 GAGCTGTGCTTCCGAGCGGAGATGATCTATCTATCTATCTATCTATCTATCTATCT 4620
Qy 4621 GCAATTAACAGCCAAAGCAGATGTCATATCAATAGAAAGATGAAATTCACCCGACACC 4680
Db 4621 GCAATTAACAGCCAAAGCAGATGTCATATCAATAGAAAGATGAAATTCACCCGACACC 4680
Qy 4681 AACTGTGTGCTGCGGAGAAACCTCTACTTGTGTAAGTCCAGCTTCCCTGACAAACAGGC 4740
Db 4681 AACTGTGTGCTGCGGAGAAACCTCTACTTGTGTAAGTCCAGCTTCCCTGACAAACAGGC 4740

QY 4741 TGGGTACCGCCCTTAGAATAGTTGTGCGAGGTGGAGAGTTTCTAGGGAAAAACAGAA 4800
DB 4741 TGGGTACCGCCCTTAGAATAGTTGTGCGAGGTGGAGAGTTTCTAGGGAAAAACAGAA 4800
QY 4801 GCTGAGTCTAAACGCTTGGAAAACCTCCCTGCTGAAAATGGAAAGTGTATGCCCTTAGAC 4860
DB 4801 GCTGAGTCTAAACGCTTGGAAAACCTCCCTGCTGAAAATGGAAAGTGTATGCCCTTAGAC 4860
QY 4861 ATGAAATGACAGCGTCCCTTCAAGTACAGAGTGTGTGTGGGCAACGAGAAAGGCTC 4920
DB 4861 ATGAAATGACAGCGTCCCTTCAAGTACAGAGTGTGTGTGGGCAACGAGAAAGGCTC 4920
QY 4921 TACGCCCTGAATGCTTGA AAAACCTCCCTAACCAGTGTCCAGGAATGGAGAGTCTTC 4980
DB 4921 TACGCCCTGAATGCTTGA AAAACCTCCCTAACCAGTGTCCAGGAATGGAGAGTCTTC 4980
QY 4981 CAAATTTATATTTATCAAGAGCTGTGAGAGAGTACTATGATAGAGAGAGAGAGGAGCA 5040
DB 4981 CAAATTTATATTTATCAAGAGCTGTGAGAGAGTACTATGATAGAGAGAGAGAGGAGCA 5040
QY 5041 CTGTGTCTTGTGAGAGTGAAGAAAGTGAACAGTCCCTGGCCAGTCCCACTGCTGCC 5100
DB 5041 CTGTGTCTTGTGAGAGTGAAGAAAGTGAACAGTCCCTGGCCAGTCCCACTGCTGCC 5100
QY 5101 CAGCCGACATCTCACCACCAATTTTGAAGCTGTCAAGGCTGCGCACTTGTGTGGGCA 5160
DB 5101 CAGCCGACATCTCACCACCAATTTTGAAGCTGTCAAGGCTGCGCACTTGTGTGGGCA 5160
QY 5161 GCGCAAGTTGAGAAAGGAGCTGTGATCTGTGACAGCATGCCAGCAAAAGTGTATCTC 5220
DB 5161 GCGCAAGTTGAGAAAGGAGCTGTGATCTGTGACAGCATGCCAGCAAAAGTGTATCTC 5220
QY 5221 CGCTACAGAGAAACCTCAGCAATACTGCATCCGGAAGAGATAGAGACCTCAGAGCCC 5280
DB 5221 CGCTACAGAGAAACCTCAGCAATACTGCATCCGGAAGAGATAGAGACCTCAGAGCCC 5280
QY 5281 TGGAGCTGTATCCACTTCAACCAATTAGAGTACTCTCATTTGGAAACCAATTAATTTACGAA 5340
DB 5281 TGGAGCTGTATCCACTTCAACCAATTAGAGTACTCTCATTTGGAAACCAATTAATTTACGAA 5340
QY 5341 ATGACATGAGAGAGTACAGCTGTGAGAAATCTGTGATAGAAATGACATTTCTTGGCA 5400
DB 5341 ATGACATGAGAGAGTACAGCTGTGAGAAATCTGTGATAGAAATGACATTTCTTGGCA 5400
QY 5401 CCGTGTGTGTGTGCGGCTCTTCCACAGCTCCCTGCTCAATCGTGCAGGTGACAGC 5460
DB 5401 CCGTGTGTGTGTGCGGCTCTTCCACAGCTCCCTGCTCAATCGTGCAGGTGACAGC 5460
QY 5461 GCAAGGAGAGAGAGTACTGCTGTGTTCACAGAAATTTGAGTGTGTGATTTCT 5520
DB 5461 GCAAGGAGAGAGAGTACTGCTGTGTTCACAGAAATTTGAGTGTGTGATTTCT 5520
QY 5521 TACGGAAGAGTACCGGACAGAGATCTCAAGTGAAGTGTCTTACCTTTGGCTTTGCC 5580
DB 5521 TACGGAAGAGTACCGGACAGAGATCTCAAGTGAAGTGTCTTACCTTTGGCTTTGCC 5580
QY 5581 TACAGAGAAACCTTATCTGTTTGTGACCACTTCAACTCACTGCAAGTATTTGAGATCCAG 5640
DB 5581 TACAGAGAAACCTTATCTGTTTGTGACCACTTCAACTCACTGCAAGTATTTGAGATCCAG 5640
QY 5641 GCAAGCTCTCAG 5700
DB 5641 GCAAGCTCTCAG 5700
QY 5701 CTGGGCTCTGCACTTTCTCAGAGAGATTTACTTGGGCTCTCATACAGATTAATTA 5760
DB 5701 CTGGGCTCTGCACTTTCTCAGAGAGATTTACTTGGGCTCTCATACAGATTAATTA 5760
QY 5761 AGGTTCAATTTGCTCAG 5820
DB 5761 AGGTTCAATTTGCTCAG 5820

QY 5821 CCGTCCACCTCCGAG 5837
DB 5821 CCGTCCACCTCCGAG 5837

RESULT 3
US-09-804-471A-1
; Sequence 1, Application US/09804471A
; Patent No. 6479269
; GENERAL INFORMATION:
; APPLICANT: WEBSTER, Marion et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; FILE REFERENCE: CL001164
; CURRENT APPLICATION NUMBER: US/09/804,471A
; INVENTION FILING DATE: 2001-03-13
; NUMBER OF SEQ. ID NOS.: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ. ID NO. 1
; LENGTH: 1515
; TYPE: DNA
; ORGANISM: Human
US-09-804-471A-1

Query Match 22.7%; Score 1398.2; DB 4; Length 1515;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1400; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ATGTGAAAGTTCAAAATATGAGAGCGCGAATCCCTTGTGATGCTGTGCTGAACCAT 60
DB 7 ATGTGAAAGTTCAAAATATGAGAGCGCGAATCCCTTGTGATGCTGTGCTGAACCAT 66
QY 61 GCCAGCGGAGCTCCAGAGTGTGCTTCCAGAGGAGAAACACCTTTATGACTCAA 120
DB 67 GCCAGCGGAGCTCCAGAGTGTGCTTCCAGAGGAGAAACACCTTTATGACTCAA 126
QY 121 CAGAGATGTCTCTCTTTCCGAGAGAGATTTAGATGCCCTTTGTCTTTGAA 180
DB 127 CAGAGATGTCTCTCTTTCCGAGAGAGATTTAGATGCCCTTTGTCTTTGAA 186
QY 181 GAATGACAGTGTGCTGTGATGAGATTAAGACAGTGAACATTTGCCGAGAT 240
DB 187 GAATGACAGTGTGCTGTGATGAGATTAAGACAGTGAACATTTGCCGAGAT 246
QY 241 TCCGACACATAGCTGATTAAGAGAGAGCTCAGGCTCCGAGAGAGAGAGAGAGAG 300
DB 247 TCCGACACATAGCTGATTAAGAGAGAGCTCAGGCTCCGAGAGAGAGAGAGAGAG 306
QY 301 AGTCTGTAGTGTGTGCTCACTTGTGATGAGTGAAGTGAAGAGAGAGAGAGAGAG 360
DB 307 AGTCTGTAGTGTGTGCTCACTTGTGATGAGTGAAGTGAAGAGAGAGAGAGAGAG 366
QY 361 GACATCTATCTATGAAAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 420
DB 367 GACATCTATCTATGAAAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 426
QY 421 TTTTGGAG 480
DB 427 TTTTGGAG 486
QY 481 CAGATGCTCTTCAAG 540
DB 487 CAGATGCTCTTCAAG 546
QY 541 GACTTGTGTCACTTTGAAATGATATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 600
DB 547 GACTTGTGTCACTTTGAAATGATATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 606
QY 601 TACTTACTGAGAGATTTTGGCTGTGACAGAGGTTATCTGATGAGATAGTGCATGCA 660
DB 607 TACTTACTGAGAGATTTTGGCTGTGACAGAGGTTATCTGATGAGATAGTGCATGCA 666


```

Db      1027 TTGTTGTGCGCCAGAAAGAGACTGAAGTTGAAGTCTTTGTCGCCATCCTTTCTTC 1086
Qy      1081 TCTAAATTTGACTGGAACAACATTCGTAATCTCTCCCCCTTCGTTCCACCTCAAG 1140
Db      1087 TCTAAATTTGACTGGAACAACATTCGTAATCTCTCCCCCTTCGTTCCACCTCAAG 1146
Qy      1141 TCTGAGATGACACCTTCCAAATTTTGAATGAACAGAGAAAGATTGCGGGTTTCATCTCT 1200
Db      1147 TCGAGAGATGACACCTTCCAAATTTTGAATGAACAGAGAAAGATTGCGGGTTTCATCTCT 1206
Qy      1201 CCGTGGCAGTGAAGCCCTTAAGCTTCTCGGGTGAAGAACTGCCGTTTGGGGTTTTCG 1260
Db      1207 CCGTGGCAGTGAAGCCCTTAAGCTTCTCGGGTGAAGAACTGCCGTTTGGGGTTTTCG 1266
Qy      1261 TACAGCAAGGCACTGGGGATTTCTTGGTATCTGAAGTCTGTGTGTCGGGTTGACTCC 1320
Db      1267 TACAGCAAGGCACTGGGGATTTCTTGGTATCTGAAGTCTGTGTGTCGGGTTGACTCC 1326
Qy      1321 CCGTCCAGACTAGCTCCATGAGAAAGAACTTCTCATCAAAAGCAAGAGCTTACAGAC 1380
Db      1327 CCGTCCAGACTAGCTCCATGAGAAAGAACTTCTCATCAAAAGCAAGAGCTTACAGAC 1386
Qy      1381 TCTCAGACACAGTGTCAAGAT 1403
Db      1387 TCTCAGACACAGTGTCAAGAT 1409

```

RESULT 5

```

US-09-774-528-419
; Sequence 419, Application US/09774528
; Patent No. 6743619
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Zhou, Ping
; APPLICANT: Goodrich, Ryle
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Ren, Feiyan
; APPLICANT: Zhang, Jie
; APPLICANT: Zhao, Qing A.
; APPLICANT: Yang, Yonghong
; APPLICANT: Xue, Aildong J.
; APPLICANT: Wehrman, Tom
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Wang, Dunrui
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. 6743619el Nucleic Acids and
; FILE REFERENCE: 802
; CURRENT APPLICATION NUMBER: US/09/774,528
; CURRENT FILING DATE: 2001-01-30
; NUMBER OF SEQ ID NOS: 441
; SOFTWARE: pl_genes Version 2.0
; SEQ ID NO 419
; LENGTH: 2162
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (364)..(2010)
US-09-774-528-419

```

```

Query Match      15.6%; Score 958.8; DB 4; Length 2162;
Best Local Similarity 99.8%; Pred. No. 6.8e-241;
Matches 960; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```

Qy      1 ATGTTGAAGTCAAAATATGAGCGCGGAATCTTTGGATCTGCTGCTGAACCAT 60
Db      364 ATGTTGAAGTCAAAATATGAGCGCGGAATCTTTGGATCTGCTGCTGAACCAT 423
Qy      61 GCCAGCGGGCTCCAGCTGAATCTGTTCTTCAGGGGAACACACCTTATGACTCA 120

```

```

Db      424 GCCAGCGGGCTCCAGCTGAATCTGTTCTTCAGGGGAACACACCTTATGACTCA 483
Qy      121 CAGCAGATGCTCTCTCTTTCCGAGAAAGGATATTAATGATGCTCTTGTCTTTGAA 180
Db      484 CAGCAGATGCTCTCTCTTTCCGAGAAAGGATATTAATGATGCTCTTGTCTTTGAA 543
Qy      181 GAATGAGTCAAGCTGCTGCTGATGAAGTTAAGACGTGAGCACTTTGTCGGAAAT 240
Db      544 GAATGAGTCAAGCTGCTGCTGATGAAGTTAAGACGTGAGCACTTTGTCGGAAAT 603
Qy      241 TCCGACACCATAGCTAGTTACAGAGAGCTCAGCTTCGGCAAGGACTTGAAGTGA 300
Db      604 TCCGACACCATAGCTAGTTACAGAGAGCTCAGCTTCGGCAAGGACTTGAAGTGA 663
Qy      301 AGTCTTGAGTGTGTGATCACTTTGCTGAAGTGCAGGTGTGAAGAGAAAGCAACGGG 360
Db      664 AGTCTTGAGTGTGTGATCACTTTGCTGAAGTGCAGGTGTGAAGAGAAAGCAACGGG 723
Qy      361 GACATCTATGCTATGAAGTATGAAGAAAGAGCTTTATGGCCCAAGAGCTTTCA 420
Db      724 GACATCTATGCTATGAAGTATGAAGAAAGAGCTTTATGGCCCAAGAGCTTTCA 783
Qy      421 TTTTGTGAGAAAGCGGAAACATTTATCTCGAAGCAAGCCCGTGAATCCCAATTA 480
Db      784 TTTTGTGAGAAAGCGGAAACATTTATCTCGAAGCAAGCCCGTGAATCCCAATTA 843
Qy      481 CAGTATGCTTTGAGGACAAATAATCACTTTATCTGATGATGAATATCAGCTGAGAG 540
Db      844 CAGTATGCTTTGAGGACAAATAATCACTTTATCTGATGATGAATATCAGCTGAGAG 903
Qy      541 GACTTGTGCTCACTTTTGAATATGATATGAGAGCAAGTTAGATGAATCACTGATACGTTT 600
Db      904 GACTTGTGCTCACTTTTGAATATGATATGAGAGCAAGTTAGATGAATCACTGATACGTTT 963
Qy      601 TACTTGAAGTGAATTTTGGCTGTTCAAGGTTCACTGATGATGAGTACGTGATCGA 660
Db      964 TACTTGAAGTGAATTTTGGCTGTTCAAGGTTCACTGATGATGAGTACGTGATCGA 1023
Qy      661 GACATCAAGGCTGAGAAACATTTCTGTTGACCGGACAGACATCAAGCTGTGATTTT 720
Db      1024 GACATCAAGGCTGAGAAACATTTCTGTTGACCGGACAGACATCAAGCTGTGATTTT 1083
Qy      721 GGATCTGCGCGGAAATGAATTTCAACAGAGTGAATGCCAACTCCGATTTGGAGC 780
Db      1084 GGATCTGCGCGGAAATGAATTTCAACAGAGTGAATGCCAACTCCGATTTGGAGC 1143
Qy      781 CCGATTTACATGAGCTCTCTGAAGTCTGATGATGAACGGGATGGAAGGACCTTAC 840
Db      1144 CCGATTTACATGAGCTCTCTGAAGTCTGATGATGAACGGGATGGAAGGACCTTAC 1203
Qy      841 GGCCTGAGCTGTGATCTGTGTGTCAGTGGGCGTGAATGCTATGATGATTTATGGAGA 900
Db      1204 GGCCTGAGCTGTGATCTGTGTGTCAGTGGGCGTGAATGCTATGATGATTTATGGAGA 1263
Qy      901 TCCCCCTTGCAGAGGAAACCTTGTGCCAGAACTTCAATTAACATTAATTTCCAGCG 960
Db      1264 TCCCCCTTGCAGAGGAAACCTTGTGCCAGAACTTCAATTAACATTAATTTCCAGCG 1323
Qy      961 TT 962
Db      1324 TT 1325

```

RESULT 6

```

US-09-916-204-1
; Sequence 1, Application US/09916204
; Patent No. 6638745
; GENERAL INFORMATION:
; APPLICANT: WEI, Ming-Hui et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; FILE REFERENCE: CL001164CIP

```

CURRENT APPLICATION NUMBER: US/09/916,204
CURRENT FILING DATE: 2001-07-24
NUMBER OF SEQ ID NOS: 6
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 1
LENGTH: 1133
TYPE: DNA
ORGANISM: Human
US-09-916-204-1

Query Match 12.3%; Score 756.4; DB 4; Length 1133;
Best Local Similarity 99.9%; Pred. No. 6.3e-188;
Matches 757; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGTTGAAGTTCAAAATATGAGAGCGCGGAATCTTTGGATGCTGCTGCTGAACCAATT 60
DB 54 ATGTTGAAGTTCAAAATATGAGAGCGCGGAATCTTTGGATGCTGCTGCTGAACCAATT 113
QY 61 GCCAGCCGGGCTCCAGGCTGAATCTGTTCTCCAGGGGAAACCACTTTATGACTCAA 120
DB 114 GCCAGCCGGGCTCCAGGCTGAATCTGTTCTCCAGGGGAAACCACTTTATGACTCAA 173
QY 121 CAGCAGATGTCCTCTTTTCCGAGAGGATATTAATGATCCCTCTTTGTTCTTTGAA 180
DB 174 CAGCAGATGTCCTCTTTTCCGAGAGGATATTAATGATCCCTCTTTGTTCTTTGAA 233
QY 181 GAATGACGTAGAGCTGCTGATGAGATTAAGCACTGAGCAACTTTGTCGGAAATAT 240
DB 234 GAATGACGTAGAGCTGCTGATGAGATTAAGCACTGAGCAACTTTGTCGGAAATAT 293
QY 241 TCCGACACCATAGCTAGTACAGAGAGCTCCAGCTTCGGCAAGAGCTTCGAAGTCAGA 300
DB 294 TCCGACACCATAGCTAGTACAGAGAGCTCCAGCTTCGGCAAGAGCTTCGAAGTCAGA 353
QY 301 AGTCTTGAGTGTGTGCTCACTTTGCTGAAGTCAAGTGTGTAAGAGAGAAAGCAACCGG 360
DB 354 AGTCTTGAGTGTGTGCTCACTTTGCTGAAGTCAAGTGTGTAAGAGAGAAAGCAACCGG 413
QY 361 GACATCTATGCTATGAAAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 420
DB 414 GACATCTATGCTATGAAAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 473
QY 421 TTTTGGAG 480
DB 474 TTTTGGAG 533
QY 481 CAGTATGCTTTTCAAG 540
DB 534 CAGTATGCTTTTCAAG 593
QY 541 GACTTGCTGCTCACTTTTGAATATGATGAGAGCAAGTATGAGAGAGAGAGAGAGAG 600
DB 594 GACTTGCTGCTCACTTTTGAATATGATGAGAGCAAGTATGAGAGAGAGAGAGAGAG 653
QY 601 TACCTAGAGCTGAGCTGATTTTGGCTGTTCAAGGCTTCATCTGATGGGATACGGATCGA 660
DB 654 TACCTAGAGCTGAGCTGATTTTGGCTGTTCAAGGCTTCATCTGATGGGATACGGATCGA 713
QY 661 GACATCAAGCTGAG 720
DB 714 GACATCAAGCTGAG 773
QY 721 GGATCTGCCGCGAGAAATGAAATCAAAAGATGTGAA 758
DB 774 GGATCTGCCGCGAGAAATGAAATCAAAAGATGTGAA 811

RESULT 7
US-10-282-048-1
Sequence 1, Application US/10282048
Patent No. 6692948
GENERAL INFORMATION:
APPLICANT: WEI, Ming-Hui et al.

TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
THEREOF
FILE REFERENCE: CL001164CIP-DIV
CURRENT APPLICATION NUMBER: US/10/282,048
CURRENT FILING DATE: 2002-10-29
NUMBER OF SEQ ID NOS: 6
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 1
LENGTH: 1133
TYPE: DNA
ORGANISM: Human
US-10-282-048-1

Query Match 12.3%; Score 756.4; DB 4; Length 1133;
Best Local Similarity 99.9%; Pred. No. 6.3e-188;
Matches 757; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGTTGAAGTTCAAAATATGAGAGCGCGGAATCTTTGGATGCTGCTGCTGAACCAATT 60
DB 54 ATGTTGAAGTTCAAAATATGAGAGCGCGGAATCTTTGGATGCTGCTGCTGAACCAATT 113
QY 61 GCCAGCCGGGCTCCAGGCTGAATCTGTTCTTCAGGGGAAACCACTTTATGACTCAA 120
DB 114 GCCAGCCGGGCTCCAGGCTGAATCTGTTCTTCAGGGGAAACCACTTTATGACTCAA 173
QY 121 CAGCAGATGTCCTCTTTTCCGAGAGGATATTAATGATCCCTCTTTGTTCTTTGAA 180
DB 174 CAGCAGATGTCCTCTTTTCCGAGAGGATATTAATGATCCCTCTTTGTTCTTTGAA 233
QY 181 GAATGACGTAGAGCTGCTGATGAGATTAAGCACTGAGCAACTTTGTCGGAAATAT 240
DB 234 GAATGACGTAGAGCTGCTGATGAGATTAAGCACTGAGCAACTTTGTCGGAAATAT 293
QY 241 TCCGACACCATAGCTAGTACAGAGAGCTCCAGCTTCGGCAAGAGCTTCGAAGTCAGA 300
DB 294 TCCGACACCATAGCTAGTACAGAGAGCTCCAGCTTCGGCAAGAGCTTCGAAGTCAGA 353
QY 301 AGTCTTGAGTGTGTGCTCACTTTGCTGAAGTCAAGTGTGTAAGAGAGAAAGCAACCGG 360
DB 354 AGTCTTGAGTGTGTGCTCACTTTGCTGAAGTCAAGTGTGTAAGAGAGAAAGCAACCGG 413
QY 361 GACATCTATGCTATGAAAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 420
DB 414 GACATCTATGCTATGAAAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 473
QY 421 TTTTGGAG 480
DB 474 TTTTGGAG 533
QY 481 CAGTATGCTTTTCAAG 540
DB 534 CAGTATGCTTTTCAAG 593
QY 541 GACTTGCTGCTCACTTTTGAATATGATGAGAGCAAGTATGAGAGAGAGAGAGAGAG 600
DB 594 GACTTGCTGCTCACTTTTGAATATGATGAGAGCAAGTATGAGAGAGAGAGAGAGAG 653
QY 601 TACCTAGAGCTGAGCTGATTTTGGCTGTTCAAGGCTTCATCTGATGGGATACGGATCGA 660
DB 654 TACCTAGAGCTGAGCTGATTTTGGCTGTTCAAGGCTTCATCTGATGGGATACGGATCGA 713
QY 661 GACATCAAGCTGAG 720
DB 714 GACATCAAGCTGAG 773
QY 721 GGATCTGCCGCGAGAAATGAAATCAAAAGATGTGAA 758
DB 774 GGATCTGCCGCGAGAAATGAAATCAAAAGATGTGAA 811

RESULT 8
US-09-016-434-513

Db 1681 CCATGGAAG---ATGGAAGAGAGATATGACCTGATATGATGCTGTGCTTTGGGG 1737
Qy 872 TGATTCCTATGAGATGATTTATGAGAGATCCCTTCGACAGGGAACCTTCGCGAA 931
Db 1738 TCTGTATGATGAATGCTTTTACGAGAAACCACTTTATGACAGATCGCTGGTGAGA 1797
Qy 932 CCTTCATTAACATTAATGATTTTCCAGCGGCTTTTGAATTTCCAG---ATACCCCAAG 988
Db 1798 CATACGAAATATCATGAAACCAAGAGAGGTTTCAGTTTCAGCCCAAGTGAATGATG 1857
Qy 989 TGACAGATGACTTTCTGATCTGATTCGAAAGCTTTGTCGCGCCAGAAAGAGAGCT-- 1046
Db 1858 TGTCTGAAATATGTAAGAGATCTTATGGAAGCTCATTTGTAAGAGAGAAATGATGCTTG 1917
Qy 1047 -----GAAGTTGAAGGCTTTTGTGCTGCCACTCTCTTCTCTTAAATGATGAAACA 1099
Db 1918 GTCAAAATGAAATGAAGACCTTTAAGAAACCCCATTTTTCAGTGAATGATTTGGGATA 1977
Qy 1100 ACATTGTAACCTCTCCCTCCCTCCCTGTCCTCCACCTCAAGTCTGACAGATGACCTTCCA 1159
Db 1978 ATATTGGAATCTGAAAGACACTTATATTCAGAAAGTTAGTACCCCAACAGATACATGGA 2037
Qy 1160 ATTATTGTAAGACAGAGAAAGATTCGTGGGTTTCATCTCTCCGTGCGAGCTGAGCCCT 1219
Db 2038 ATTATTGTAAGATGATGATTTGTTTAAATAATCTGAAAGCATGCCCCCAACACATA 2097
Qy 1220 CAGGCTCTCGGCTGGAAGAACTGCCCTTTGTGGGGCTTTGTGTACA 1264
Db 2098 CTGCAATTTCTGGCCACCATCTGCAATTTGTTGTTTACATATA 2142

RESULT 10

US-08-422-699A-12
Sequence 12, Application US/08422699A
Patent No. 5955265
GENERAL INFORMATION:
APPLICANT: Brook, J. David
APPLICANT: Housman, David E.
APPLICANT: Shaw, Duncan J.
APPLICANT: Hartley, Keith J.
APPLICANT: Johnson, Keith J.
TITLE OF INVENTION: DNA SEQUENCE ENCODING THE MYOTONIC
TITLE OF INVENTION: DYSTROPHY GENE AND USES THEREOF
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
STREET: Two Millitia Drive
City: Lexington
STATE: Massachusetts
COUNTRY: US
ZIP: 02713
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION NUMBER: US/08/422,699A
APPLICATION NUMBER: US/08/422,699A
CLASSIFICATION:
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: 08/422,706
FILING DATE:
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 08/023,612
FILING DATE: 26-FEB-1993
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 07/839,255
FILING DATE: 20-FEB-1992
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: PCT/US93/01545
FILING DATE: 19-FEB-1993
PRIORITY APPLICATION DATA:

APPLICATION NUMBER: PCT/GB93/00253
FILING DATE: 05-FEB-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB9202485.0
FILING DATE: 06-FEB-1992
ATTORNEY/AGENT INFORMATION:
NAME: Granahan, Patricia
REGISTRATION NUMBER: 32,227
REFERENCE/DOCKET NUMBER: MIT-5830A2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-861-6240
TELEFAX: 617-861-9540
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 2726 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
US-08-422-699A-12

Query Match 3.7%; Score 227.2; DB 2; Length 2726;
Best Local Similarity 54.1%; Pred. No. 9.3e-49;
Matches 535; Conservative 0; Mismatches 438; Indels 15; Gaps 3;

Qy 194 CTGCTGATGAAAGATTAGACAGTGAACATTTGTCCGAAAGTATTCGACACATAG 253
Db 243 CCGAATCTGGCCCGACGACCAAGTACGTGCGGCACTTCTTGCAGTGGCGAGCCATCGTGG 302
Qy 254 CTGAGTTACAGAGCTCCAGCCCTTCGCAAAAGACTTCGAAGTCAGAGCTTGTAGT 313
Db 303 TGAGGCTTAAGAGAGCTCCAGCTGCAAGAGTGAAGAGAGAGAGAGAGAGAGAGAGAG 362
Qy 314 GTGCTCACTTGTCTGAAGTGCAGGTGTGAAGAGAAAGCAACCGGAGCATCTATGCTA 373
Db 363 GCGGGCGGTTACGCGAGGTAGAGGTAGTGAAGTGAAGAGAGAGAGAGAGAGAGAGAGAG 422
Qy 374 TGAAGTGAAG 433
Db 423 TGAAGTGAAG 482
Qy 434 AGCGAACAATATATTCGAAAGCAAGCCGCTGATCCCAATTAAGTACGTAAGCTTTC 493
Db 483 AGAGGACGTGTGTGTAATGGAGACCGGCGGTGATACGAGCTGCACTTCGCTTCC 542
Qy 494 AGGCAAAATATGACCTTTATCTGTCATGAAATATCAGCTGAGAGGAGAGAGAGAGAGAG 553
Db 543 AGGATAGAGAACTACCTGTACCTGTGATAGAGATATACGTGGCGGAGAGAGAGAGAGAG 602
Qy 554 TTTTGAATGATATGAG 613
Db 603 TGCTGAGCAAGTTTGGGAGCGGATTCGCGCGAGATGCGGCTTCTTACCTGGCGGAGA 662
Qy 614 TGATTTTGGCTGTCAACGCTTCAATCTGAAGAGATGAGATGAGATGAGATGAGATGAGAT 673
Db 663 TTGTATGAGCAATGAGCTGTGAGACCGGCTTGGCTGACGAGAGAGAGAGAGAGAGAGAGAG 722
Qy 674 AGAATCTTCGTTGACGAG 733
Db 723 ACAACATCTGTGAG 782
Qy 734 AATGATTAACAAAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGAT 793
Db 783 AGCTGCGGAGAGATGAG 842
Qy 794 CTTCGAGAGTGTGA---CTGTGATGAAAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 850
Db 843 CCGCGAGATCTGTGAG 902
Qy 851 GTGAGTGTGCTCAGTGGCGGTGATTTGCTATGATGATGATGATGATGATGATGATGATGAT 910
Db 903 GTGATGTGTGGGCGGTGTGATTTGCTATGATGATGATGATGATGATGATGATGATGATGAT 962
Qy 911 CAGAGGAACTCTGCGAGAACTTCAATTAACATTATGATGATTTCCAGGCGGTTTGAAT 970

Db 963 ACGGGATTCCACGGCCGAGCCTATGCAAGATCTCAACAGAGACCTCTCTC 1022
Qy 971 TTCAGATGACCCCAAGTGAAGT---GACTTCTTGAATCTGATTCAAAGCTTGTGT 1027
Db 1023 TGCCGCTGCTGAGCAAGAGGGTCCCTGAGAGGCTCCAGACTTATTCAGGGCTTGTCT 1082
Qy 1028 GCGGCCGAGAAAGAGAGCTGAAGTTTGAAGTCTTTTCTCTG-----CATCTTTCT 1078
Db 1083 GTCCCCCGAGACACAGCTGGGCGGGGTGGAGCAGGACTTCGAGACATCCCTTCT 1142
Qy 1079 TCTTTAAATGATGATGAAACAATTTGTAATCTCTCTCCCTTCTGTTCCACCTCTCA 1138
Db 1143 TCTTTGCTCTGACTGGAGTGTCTCCGGACAGCTGCCCCCTTTACACCGGATTTCTG 1202
Qy 1139 AGCTGACATGACACCTCCATTTTGA 1166
Db 1203 AAGTCCACCGACATGCACTTTCGA 1230

RESULT 11

US-08-422-706B-12

Sequence 12. Application US/08422706B
Patent No. 5977333

GENERAL INFORMATION:

APPLICANT: Brook, J. David

APPLICANT: Houman, David E.

APPLICANT: Shaw, Duncan J.

APPLICANT: Harley, Helen G.

APPLICANT: Johnson, Keith J.

TITLE OF INVENTION: DNA SEQUENCE ENCODING THE MOTONIC

NUMBER OF SEQUENCES: 14

CORRESPONDENCE ADDRESSES: 14

ADDRESSES: Hamilton, Brook, Smith & Reynolds, P.C.

STREET: Two Militia Drive

CITY: Lexington

STATE: Massachusetts

COUNTRY: US

ZIP: 02713

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/422,706B

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/284,543

FILING DATE: 08-AUG-1994

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/023,612

FILING DATE: 26-FEB-1993

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/839,255

FILING DATE: 20-FEB-1992

PRIOR APPLICATION DATA:

APPLICATION NUMBER: PCT/US93/01545

FILING DATE: 19-FEB-1993

PRIOR APPLICATION DATA:

APPLICATION NUMBER: PCT/GB93/00253

FILING DATE: 05-FEB-1993

PRIOR APPLICATION DATA:

APPLICATION NUMBER: GB9202485.0

FILING DATE: 06-FEB-1992

ATTORNEY/AGENT INFORMATION:

NAME: Granahan, Patricia

REGISTRATION NUMBER: 32,227

REFERENCE/DOCKET NUMBER: MIT-5830A2

TELECOMMUNICATION INFORMATION:

TELEPHONE: 617-861-6240

TELEFAX: 617-861-9540
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2726 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
US-08-422-706B-12

Query Match 3.7%; Score 227.2; DB 2; Length 2726;
Best Local Similarity 54.1%; Pred. No. 9,3e-49;
Matches 535; Conservative 0; Mismatches 438; Indels 15; Gaps 3;

Qy 194 CTGCTGATGAAAGATTAAACAGCTGAGCAATTTGTCGGAAATATTCGACACCATAG 253
Db 243 CCGAACTGGCCCGAGACAAAGTACGTGGCCGACCTTCTGCAATGGGCGGAGCCCATCGGG 302
Qy 254 CTGAGTTACAGAGCTTCACGCTTCGCGCAAGAGCTTGAAGTGAAGTCTTAGGTT 313
Db 303 TGAAGCTTAAAGAGTCCGACTGCAAGAGGACACTTGAAGATTCTGAAGTATCGGAC 362
Qy 314 GTGGTCACTTGTCTGAAGTCCAGGTGTTAAGAGAAAGCAACCGGAGACATCTATGCTA 373
Db 363 GCGGGCGCTTCAGGAGGTAGCGGTAGTGAAGTGAAGCAAGCGGCCAGGTGATGCCA 422
Qy 374 TGAAGTGAAGAAAGAGAGCTTATTTGCCCCAGAGAGTTTCATTTTGAAGAG 433
Db 423 TGAAGATCATGAACAAAGTGGAGACATGCTGAAGAGGGCGAGGTGCTGCTCGTGAAG 482
Qy 434 AGCGGAACATTTATCTCTGAAGCAACAGCCCGTGAATCCCAATTACAGTATGCTTTC 493
Db 483 AGAGGAGCGTGTGTGAATGGGAGCCGGGTGATGACGACACTTGCCTTCC 542
Qy 494 AGGACAAAATCACTTTATCTGTGTCATGGAATTCAGCTGAGGAGACTTGTCTCAC 553
Db 543 AGATGAAACTTACTTACTTACTGTCATGAGATTTACGTGGGCGGAGACTGCTGACAC 602
Qy 554 TTTGAAATAGATAGAGAGACCAATAGTAAACCTGATACATTTTACCTAGCTGAGC 613
Db 603 TGCTGACAAAGTTTGGGAGCGGATTCGGCCGAGATGGCGCTTCTACCTGCGGAGA 662
Qy 614 TGATTTGGCTGTTCAACAGCTTATCTGATGGATACGTGATCGAGACATCAAGCTG 673
Db 663 TTGTCAATGGCCATAGACTCTGCTGACCGGCTTGGCTACGTGCAAGGACATCAACCG 722
Qy 674 AGAACAATTCGTTGACCGGACAGGACATCAAGCTGTGATTTGATCTGCGGGA 733
Db 723 ACAACATCTGCTGAGCCGCTGTGCAACATCGCCTGCGGACTTGGCTTTCCTCA 782
Qy 734 AATGAATTCAAACAAAGATGATGCAAACTCCGATTTGGAGCCCGAGATTACATGG 793
Db 783 AGCTGCGGGAGATGAAAGGTGGCGTGGCTGTGGTGGTGGGACCCAGACTACCTGT 842
Qy 794 CTCTGAAGTCTGA---CTGTGATGAACGGGATGGAAGCACTACGCGCTGAGCT 850
Db 843 CCCCCGAGATCTGCAAGGCTGTGGCGGCTGTGGAGACAGGCTACCGGCGCAAGT 902
Qy 851 GTGACTGTGTGATGAGGCGGTGATTTGCTATGAGATTTTATGGAAGATCCCTTTCG 910
Db 903 GTGACTGTGTGCGCTGTGGGTGATTTGCTATGAAATGTTCTATGAGGACAGCGCTTCT 962
Qy 911 CAGAGGAACTCTGCGAGAACTTCAATTAATTAATTAATTTCCAGCGTTTGTGAAT 970
Db 963 ACGGGATTCCACGGCGGAGACCTATGCAAGATCGTCACTAAGAGAGCACTCTCTC 1022
Qy 971 TTCAGATGACCCCAAGTGAAGT---GACTTCTTGAATCTGATTCAAAGCTTGTGT 1027
Db 1023 TGCCGCTGCTGAGCAAGAGGGTCCCTGAGAGGCTCCAGACTTATTCAGGGCTTGTCT 1082
Qy 1028 GCGGCCGAGAAAGAGAGCTGAAGTTTGAAGTCTTTTCTCTG-----CATCTTTCT 1078
Db 1083 GTCCCCCGAGACACAGCTGGGCGGGGTGGAGCAGGACTTCGAGACATCCCTTCT 1142

1079 TCTTAAATGAGTGAACAACATGCTCTCTCCCTTCCGACCTCA 1138
Db 1143 TCTTTGGCTGACTGGATGTCTCCGGACAGCGTCCCTTACACCGGATTTGG 1202
Qy 1139 AGTCTGACGATGACACTCCAAATTTGA 1166
Db 1203 AAGGTGCCACCGACATGCAACTTCGA 1230

RESULT 12
US-08-422-699A-8
Sequence 8, Application US/08422699A
Patent No. 5955265
GENERAL INFORMATION:
APPLICANT: Brook, J. David
APPLICANT: Houseman, David E.
APPLICANT: Shaw, Duncan J.
APPLICANT: Harley, Helen G.
APPLICANT: Johnson, Keith J.
TITLE OF INVENTION: DNA SEQUENCE ENCODING THE MYOTONIC
TITLE OF INVENTION: DYSTROPHY GENE AND USES THEREOF
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
STREET: Two Militia Drive
CITY: Lexington
STATE: Massachusetts
COUNTRY: US
ZIP: 02713
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/422,699A
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/422,706
FILING DATE:
APPLICATION DATA:
APPLICATION NUMBER: US 08/023,612
FILING DATE: 26-FEB-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/839,255
FILING DATE: 20-FEB-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US93/01545
FILING DATE: 19-FEB-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/GB93/00253
FILING DATE: 05-FEB-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB9202485.0
FILING DATE: 06-FEB-1992
ATTORNEY/AGENT INFORMATION:
NAME: Granahan, Patricia
REGISTRATION NUMBER: 32,227
REFERENCE/DOCKET NUMBER: MIT-5830A2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-861-6240
TELEFAX: 617-861-9540
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 2511 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1746

US-08-422-699A-8
Query Match 3.6%; Score 220.4; DB 2; Length 2511;
Best Local Similarity 54.6%; Pred. No. 5,4e-47;
Matches 512; Conservative 0; Mismatches 411; Indels 15; Gaps 3;

Qy 244 GACACATAGCTGATACAGAGCTTCAGCCTTCGGCAAAAGACCTTCGAAGTCAGAGT 303
Db 43 GCCATCGTGTGAGGCTTAAAGAGTCCGACTGAGAGGACGACCTTCGAGATTCGTAAG 102
Qy 304 CTGTAGTGTGTGCTCACTTGTGGAAGTGCAGGTGTGTAGAGAAAGCAACCGGGGAC 363
Db 103 GTGATCGACGCGGGGCTTACAGGAGTACCGGTAGTAAGTGAAGCAAGCGGCGAG 162
Qy 364 ATCTAGCTATGAAAGTGAAGAAAGAGCTTTATTTGGCCGAGAGAGCTTCAATTT 423
Db 163 GTGATAGCCATGAAGATCAAGAACAGTGGGAATGTGAAGAGGGGCGAGGTCTGTC 222
Qy 424 TTGAGGAAGAGCGGAACATATTATCTCGAAGCAAGCCGTGATCCCAATTACAG 483
Db 223 TTCCGTAGAGAGAGGACGTTGTGAATGGGACCGGGCGTGATCAGCAGCTGCAC 282
Qy 484 TATGCTTTCAGAGCAAAATCACTTTATCTGTGATGAATATACGCTGAGGGAGAC 543
Db 283 TTGCTTCAGAGATGAAGATCACTGTACTGTGTATGATGATTAAGTGGGCGGAGAC 342
Qy 544 TTGCTGTCACTTTGAATAGATATAGAGACCAATTAAGTAAACCTGATACAGTTTAC 603
Db 343 CTGCTGACACTGTGAGCAAGTTTGGGAGCGGATTCGGCCGAGATGGGCGCTTAC 402
Qy 604 CTAGCTGAGCTGATTTTGGCTGTTCACAGGTTCACTGTATGATGGATACGTGACAGAC 663
Db 403 CTGGCGAGATGTCTATGGCCATAGACTGTGTCACCGGCTTGTGCTACGTCACAGGAGAC 462
Qy 664 ATCAAGCTGAGAACTTCTCGTTGACCGCACAGACATCAAGCTGTGTGATTTTGA 723
Db 463 ATTAACCCGACAACTCTGCTGAGCCGTGTGACCACTCGCTGGCGGACTTCGGC 522
Qy 724 TCTGCGGCAAAATGATTCAAACAGATGTGAATGCAAACTCCGATTGGGACCCCA 783
Db 523 TCTTGCTCAAGCTGGGGCAGATGGAAGGTGCGGTCTGTGTGCTGTGGCACTCCA 582
Qy 784 GATTACATGCTCTCTGA---AGTCTGACTGTGATGAACGGGATGAAAAGGACCTTAC 840
Db 583 GACTACTGTCCCGGAGATCTCTGACGCTGTGGCGGTGGGCTTGGACAGGACACTAC 642
Qy 841 GGCTGTGACTGTGACTGTGTGCAAGTGGGCGGTGATTCCTATGAAGTATTAATGGAGA 900
Db 643 GGCCCGAGTGTGACTGTGGGCGCTGGGTGTATTCGCTATGAATGTCTATAGGGCAG 702
Qy 901 TCCCTCTTCAGAGAGGAACCTCTGCAGAACCTTCAATTAATGAATTTCCAGCGG 960
Db 703 AGCCCTTCTAGCGGATTTCCACGGCGGAGACTATAGCAAGTGTCCACTCAAGAGAG 762
Qy 961 TTTTGAATTTTCAGATGACCCCAAGTAGACAGT---GACTTCTTGATCTGATTCAA 1017
Db 763 CACTCTCTCTCCGCTGTGTGACGAAGGGGTCTCTGAGAGAGCTTCGACACTTCAATTGAC 822
Qy 1018 AACTTGTGTGTGGGCGCAAAAGAGACTGAAGTTGAAGTCTTGTCTGCGC----- 1068
Db 823 CGGTGCTGTGTGCCCCCGGAGACAGCGCTGGGCGGGGTGGAGCAGGCGCACTTCCGGACA 882
Qy 1069 CATCCCTTCTCTCTAAATTAATGACTGGAACAACATTCGPACTCTCCCTTCTGTT 1128
Db 883 CATCCCTTCTTCTTTGGGCTGTGACTGGATGTCTCTCCGGACAGAGGTGCCCTTCTTAC 942
Qy 1129 CCGACCTTCAAGTGTGAGATGACACTTCAATTTTGA 1166
Db 943 CCGATTTTGAAGGTGCCACCGACACATGCAACTTGA 980

RESULT 13
US-08-422-706B-8

Sequence 8, Application US/08422706B
Patent No. 5977333
GENERAL INFORMATION:
APPLICANT: Brook, J. David
APPLICANT: Houman, David E.
APPLICANT: Shaw, Duncan J.
APPLICANT: Hatley, Helen G.
APPLICANT: Johnson, Keith J.
TITLE OF INVENTION: DNA SEQUENCE ENCODING THE MYOTONIC
TITLE OF INVENTION: DYSTROPHY GENE AND USES THEREOF
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
STREET: Two Millitia Drive
CITY: Lexington
STATE: Massachusetts
COUNTRY: US
ZIP: 02713
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/422,706B
FILING DATE: 14-APR-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/284,543
FILING DATE: 08-AUG-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/023,612
FILING DATE: 26-FEB-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/839,255
FILING DATE: 20-FEB-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US93/01545
FILING DATE: 19-FEB-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/GB93/00253
FILING DATE: 05-FEB-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB9202485.0
FILING DATE: 06-FEB-1992
ATTORNEY/AGENT INFORMATION:
NAME: Granahan, Patricia
REGISTRATION NUMBER: 32,227
REFERENCE/DOCKET NUMBER: MIT-5830A2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-861-6240
TELEFAX: 617-861-9540
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 2511 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1746
US-08-422-706B-8

Query Match 3.6%; Score 220.4; DB 2; Length 2511;
Best Local Similarity 54.6%; Pred. No. 5.4e-47;
Matches 512; Conservative 0; Mismatches 411; Indels 15; Gaps 3;

244 GACACATAGCTGAGTACAGAGCTCCAGCCTTCGGCAAGCACTTCGAGTCAGAGT 303
Db 43 GCCATCTGTGAGAGCTTAAGAGAGTCCGATCGACAGAGGAGCACTTCGAGATTCGAG 102
304 CTGTAGGTTGTGCTCACTTGTCTGAAGTGCAGGTGTGTAAGAGAGAAAGCAAGCGGAGAC 363

Db 103 GTGATCGAGCGGGGGCTTCAGGAGGTAGCGGTAGTGAAGATGAACGACGGGCCAG 162
Qy 364 ATCTATGCTATGAAGATGAAGAGAGAGCTTTATTTGGCCGACAGAGCTTTCAATT 423
Db 163 GTGATGCCATGAAGATCAAGACATGAGAGACATGCTGAAGAGAGGGGAGGTCTGTC 222
Qy 424 TTGAGAGAGCGGAAATATTAATTCGAAAGACAAAGCCGTGATCCCAATTACAG 483
Db 223 TTCCGTAGAGAGAGGAGCGTGTGTGTGAATGGGGACCGGGGTGTGATACGACCTGC 282
Qy 484 TATGCCCTTTCAGAGCAAAATACCTTATCTGTGTCATGAAATATACCTCGAGGGGAC 543
Db 283 TTGGCTTTCAGAGTGAACCTACTGTACTGTGTCAATGAGATTAATCTGGGGGAGC 342
Qy 544 TTGCTGTCACTTTGAATATGATATGAGACCAAGTTAGATGAAAACCTGATCACTTTAC 603
Db 343 CTGCTGACATCTGCTGACACAGTTTGGGGAGCGATTCGGCCGAGATGGCCGCTTCTAC 402
Qy 604 CTAGCTGAGCTGATTTTGGCTGTTCAAGCGTTCATCTGATGGAGATACGTGATGAGAC 663
Db 403 CTGGCGAGATTTGTCATGGCCATAGACTCGGTGACAGCGGCTTGGCTACGTGACAGAG 462
Qy 664 ATCAAGCTGAGAACTTTCTGTTGACCGCAGAGAGACATCAAGCTGTGGATTTTGA 723
Db 463 ATCAAAACCGACAACTCTGCTGAGACCGTGTGGCCACATCCGCTGGCCGACTTGGC 522
Qy 724 TCGCCGCGAAAATGAATTCMAACAGATGATGATCCAAACTCCGATTTGGAGCCCA 783
Db 523 TCTTGCTCAAGCTGCGGCGAGATGAGACGGTGGGTGCTGTGCTGTGGGACACCCCA 582
Qy 784 GATTACATGAGCTCTCTGA--AGTGTGACTGTGATGAGACGGGATGAGAAAGGACCTTAC 840
Db 583 GACTACCTGTCCCGAGATCTGTGCAAGCTGTGGCGGTGGGCTTGGGACAGGACCTTAC 642
Qy 841 GGCTGAGCTGTGACTGTGTGTGATGAGTGGGCTGATTTGCTATGAGATTTATGGAGA 900
Db 643 GGGGCCAGATGTGACTGTGGGGCTGTGGTGTATTCCTATGAAATGTTCTATGGGCGAG 702
Qy 901 TCCGCTTTCGAGAGGAGACCTGTGCGAGAACCTTCAATTAATTGAAATTTTCACGCG 960
Db 703 AGCCCTTCTTACCGGATTTCCAGCGCGAGACCTTATGGCAAGATCTTCACTACAGAG 762
Qy 961 TTTTGAATTTTCAGATGAGACCCCAAGTGAAGT--GACTTCTTGATCTGATTCAA 1017
Db 763 CACTCTCTGCGCGGTGTGAGAGAGGGGTCCCTGAGAGGCTCGAATCTTATTTCAG 822
Qy 1018 AGCTTGTGTGCGCCAGAAAGAGACTGAATTTGAAGTCTTTGCTG----- 1068
Db 823 CGGTGTGTGTCCCGGAGACACGGCTGGGCGGGGTGGAGAGCGACTTCCGAGCA 882
Qy 1069 CATCCCTTCTCTCTAATAATTGACTGGAACAATGTTACTCTCTCCCTCGTGT 1128
Db 883 CATCCCTTCTCTTGTGCTGACTGAGATGTCTCGGGACAGCGTCCCTCTTACA 942
Qy 1129 CCCACCTCAAGTCTGAGATGACACCTCCCAATTTGA 1166
Db 943 CCGGATTTGAGAGTGCACCGACCATGCACTTGA 980

RESULT 14
US-08-484-044-11
Sequence 11, Application US/08484044
Patent No. 5552282
GENERAL INFORMATION:
APPLICANT: Caskey, C. T.
APPLICANT: Fu, Ying-Hui
APPLICANT: Friedman, David L.
APPLICANT: Pizutti, Antonio
APPLICANT: Fenwick, Raymond G.
TITLE OF INVENTION: Diagnosis of Myotonic Muscular Dystrophy
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:

ADDRESSEE: Fulbright & Jaworski, L.L.P.
STREET: 1301 McKinney, Suite 5100
CITY: Houston
STATE: Texas
COUNTRY: U.S.A.
ZIP: 77010-3095
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/484,044
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/019,940
FILING DATE: 19-FEB-1993
ATTORNEY/AGENT INFORMATION:
NAME: Paul, Thomas D.
REGISTRATION NUMBER: 32,714
REFERENCE/DOCKET NUMBER: D-5443
TELECOMMUNICATION INFORMATION:
TELEPHONE: 713/651-5325
TELEFAX: 713/651-5246
TELEX: 762829
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 3182 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-484-044-11

Query Match 3.6%; Score 220; DB 1; Length 3182;
Best Local Similarity 54.6%; Pred. No. 7.9e-47;

Matches 511; Conservative 0; Mismatches 410; Indels 15; Gaps 3;

246 CACCATAGCTGAGTTACAGAGCTCCAGCCTTCGGCAAGAACTTCAGATCAGAACTCT 305
743 CATCGTGTGAGGCTTAAAGAGGTCCGACTCGACAGAGGAGAGCTTCAGATTCTGAAGT 802
306 TGTAGTTGTGTCATCTTCTCTGAAGTGCAGGTGTGTAAGAGAAACCAACCGGGGACAT 365
803 GATCGGACCGGGGCGTTCAGCGAGGTAGCGTAGTGAATGAAAGACGAGCGGGACAGT 862
366 CTATGCTATGAAGTGTGAAGAAAGGCTTATTTGGCCCGAGAGAGATTTCAATTTT 425
863 GTATGCTATGAAGTGTGAAGAAAGGCTTATTTGGCCCGAGAGAGATTTCAATTTT 922
426 TGAGGAAGACCGGACATATTTATCTCGAAGCAACAGCCCGTGTATCCCCAATTACAGTA 485
923 CCGTGAAGAGGAGCGTGTGTGTAATGGGACCGCGGTGTATACGAGAGTGCATT 982
486 TGGCTTTGAGCAAAATATCACTTTATCTGTCATGGAATATCAGCTGTGAGGGGACTT 545
983 CGCTTTCAGGAGAACTTACCTGTACCTGTGTCATGAGTATTAAGTGTGGCGGGACCT 1042
546 GCTGTCACTTTGAATGAAATGATGAGCAAGTATGATGAAATGAAATGAAATGAAATGAA 605
1043 GCTGTCACTGTGAGCAAGTATTTGGGAGCGGATTTCCGCGGAGATGCGCGGCTTCACT 1102
606 AGCTGAGCTGATTTTGGCTGTTTCAACAGCGTTTCATGTGATGAGTATGAGTATGAGTAT 665
1103 GGGGAGATTTGTCATGCGCATAGACTCGGTGACCGGGCTTGGTACGTGACAGGAGACT 1162
666 CAAAGCTGAGAACTTCTGTTGACCGCAACAGGACATCAAGCTGTGATTTTGGAT 725
1163 CAAACCCGCAACATCTCTGGAACCGCTGTGGCCATCCGCTGTGCGGACCTTGGGCTC 1222
726 TGGCGGAAATGAAATTTCAACAGTGTGTAATGCAAACTCCGATTTGGGACCCGACA 785

DB 1223 TTGCTCAAGCTGCGGGCAGATGGAACGGTCCGTGCTGTGCTGTGGACCCACCA 1282
QY 786 TTACATGAGCTCTCTGA---AGTGTGATGTGATGAAACGGGATGGAAGGACCTACGG 842
DB 1283 CTACCTGTCCCCCGAGATCTCTGACAGGCTGTGGCGGCTGTGGACAGGACCTACGG 1342
QY 843 CTGAGCTGTGATGTGTTGTCAGTGGGCGGTGATTTGCTTATGATGATTTATGGAAGATC 902
DB 1343 GCCGAGTGTGATGTGTTGTTGGGCGCTGGGTGTATTCCTATGAAATGTTCTATGGGACAC 1402
QY 903 CCCCTTGAGAGGAGAACTCTGCGCAACCTTCATTAATATGATTTGCAAGCGCTT 962
DB 1403 GCCCTTCTAGCGGATTTCCACGGCGAGACCTTATGCAAGATGCTCACTACAGAGACA 1462
QY 963 TTGAAATTTCCAGATGATGACCCCAAGTGAACAGT---GACTTGTGATGTATTCAAAG 1019
DB 1463 CCTCTCTCTGCTGCTGTGTGATGACAGAGGGTCTCTGAGAGAGGCTGAGACTTACAGCG 1522
QY 1020 CTGTTGTGCGGCGCAGAAAGAGACTGAAAGTTGAAAGTCTTTGCTGC-----CA 1070
DB 1523 GTTGTGTGTCTCCCGGAGACAGGCTGGGCGGGGTGAGACAGGCACTTCCGACACA 1582
QY 1071 TCTCTTCTCTCTAAATTTGATGGAACCAACTTCTCTCTCTCTCTCTCTCTCTCTCTCT 1130
DB 1583 TCCCTTCT 1642
QY 1131 CACCTCAAGTGTGAGATGACAGTACACTTCAATTTTGA 1166
DB 1643 GATTTGAAAGTGTGCCACCGACACATGCAACTTGA 1678

RESULT 15
US-09-804-471A-3
; Sequence 3, Application US/09804471A
; Patent No. 6479269

GENERAL INFORMATION:
; APPLICANT: WEBSTER, Marion et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001164
; CURRENT APPLICATION NUMBER: US/09/804,471A
; CURRENT FILING DATE: 2001-03-13
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 174493
; TYPE: DNA
; ORGANISM: Human
; NAME/KEY: misc feature
; LOCATION: (1)..(174493)
; OTHER INFORMATION: n = A, T, C or G
US-09-804-471A-3

Query Match 3.3%; Score 205; DB 4; Length 174493;
Best Local Similarity 100.0%; Pred. No. 7.6e-42;

Matches 205; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 753 GGTGAATGCAAACTCCGATTTGGGACCCAGATTAACATGCTCTCTGAAGTCTGACTGT 812
DB 130289 GGTGAATGCAAACTCCGATTTGGGACCCAGATTAACATGCTCTCTGAAGTCTGACTGT 130348
QY 813 GATGAACGGGAGATGGAAGAAAGCACTACGCGCTGTGACTGTGACTGTGACTGTGACTGT 872
DB 130349 GATGAACGGGAGATGGAAGAAAGCACTACGCGCTGTGACTGTGACTGTGACTGTGACTGT 130408
QY 873 GATTCCTATGAGATGATTTATGAGAGATCCCTTTGCGAGAGGAACTCTGCGCAAGAC 932
DB 130409 GATTCCTATGAGATGATTTATGAGAGATCCCTTTGCGAGAGGAACTCTGCGCAAGAC 130468
QY 933 CTTCAATAACATTATGAATTTTCAG 957

Wed May 18 09:47:51 2005

us-10-791-666-1.rn1

Page 17

Db 130469 CTTCAATPACATTATGAATTCCAG 130493

Search completed: May 16, 2005, 18:37:09
Job time : 660 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 16, 2005, 15:06:39 ; Search time 2152 Seconds
(without alignments)
1753.063 Million cell updates/sec

Title: US-10-791-666-1
Perfect score: 6165
Sequence: 1 atgttgaaagttcaaatatg.....gggaccagctctcagatataa 6165

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 5662332 seqs, 3060109652 residues
Total number of hits satisfying chosen parameters: 11324664

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA:*

1:	/cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2:	/cgn2_6/ptodata/1/pubpna/PCTUS_NEW_PUB.seq:*
3:	/cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4:	/cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5:	/cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6:	/cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
7:	/cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8:	/cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9:	/cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
10:	/cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
11:	/cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
12:	/cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
13:	/cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
14:	/cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
15:	/cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
16:	/cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:*
17:	/cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:*
18:	/cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq:*
19:	/cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
20:	/cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
21:	/cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
22:	/cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	6165	100.0	6165 13	US-10-028-946-1 Sequence 1, Appl1
2	6165	100.0	6165 18	US-10-791-666-1 Sequence 1, Appl1
3	6165	99.9	6298 17	US-10-415-011-43 Sequence 43, Appl1
4	6144	99.7	8656 18	US-10-618-941-1 Sequence 1, Appl1
5	6121.8	99.3	6189 11	US-09-964-956-10 Sequence 10, Appl1
6	6121.8	99.3	6189 17	US-10-262-511-1 Sequence 1, Appl1
7	6105.4	99.0	6201 11	US-09-964-956-8 Sequence 8, Appl1
8	6105.4	99.0	6201 17	US-10-262-511-13 Sequence 13, Appl1
9	5837	94.7	5877 13	US-10-028-946-3 Sequence 3, Appl1
10	5837	94.7	5877 18	US-10-791-666-3 Sequence 3, Appl1
11	5661.4	91.8	6159 13	US-10-017-216-3 Sequence 3, Appl1

12	5661.4	91.8	6162 16	US-10-325-430-11 Sequence 11, Appl1
13	5661.4	91.8	6574 16	US-10-017-216-1 Sequence 1, Appl1
14	5661.4	91.8	6574 13	US-10-325-430-10 Sequence 10, Appl1
15	5661.4	91.8	6574 18	US-10-757-262-51 Sequence 51, Appl1
16	2928.2	47.5	3131 17	US-10-276-774-137 Sequence 137, Appl1
17	2682.8	43.5	2896 18	US-10-357-930-30150 Sequence 30150, A
18	2476.6	40.2	2497 17	US-10-262-511-5 Sequence 5, Appl1
19	2421.6	39.3	2542 17	US-10-262-511-7 Sequence 7, Appl1
20	1847.8	30.0	1870 17	US-10-262-511-9 Sequence 9, Appl1
21	1847.4	30.0	1870 17	US-10-262-511-3 Sequence 3, Appl1
22	1791.2	29.1	1915 17	US-10-262-511-11 Sequence 11, Appl1
23	1398.8	22.7	2066 17	US-10-311-034-33 Sequence 33, Appl1
24	1398.2	22.7	1535 9	US-09-804-471A-1 Sequence 1, Appl1
25	1398.2	22.7	1535 14	US-10-238-709-1 Sequence 1, Appl1
26	1398.2	22.7	1515 17	US-10-724-594-1 Sequence 419, Appl1
27	958.8	15.6	2162 17	US-10-120-988-419 Sequence 1, Appl1
28	957	15.5	957 17	US-10-412-897-1 Sequence 1, Appl1
29	756.4	12.3	1133 14	US-10-282-046-1 Sequence 1, Appl1
30	701.4	11.4	995 17	US-10-276-774-255 Sequence 255, Appl1
31	443.4	7.2	446 18	US-10-357-930-15814 Sequence 15814, A
32	437.8	7.1	2870 17	US-10-425-114-26241 Sequence 26241, A
33	436	7.1	485 18	US-10-357-930-45643 Sequence 45643, A
34	350.4	5.7	354 13	US-10-040-739-9 Sequence 9, Appl1
35	258	4.2	258 17	US-10-305-720-513 Sequence 513, Appl1
36	246.8	4.0	1530 17	US-10-415-011-24 Sequence 24, Appl1
37	246.8	4.0	4698 18	US-10-702-496-1 Sequence 1, Appl1
38	246.8	4.0	5372 17	US-10-333-314-40 Sequence 40, Appl1
39	246.8	4.0	5373 18	US-10-737-450-51 Sequence 51, Appl1
40	246.8	4.0	5438 18	US-10-618-941-2 Sequence 2, Appl1
41	242	3.9	2755 19	US-10-764-420-1653 Sequence 1653, Ap
42	241.6	3.9	6335 17	US-10-388-934-85 Sequence 85, Appl1
43	241.6	3.9	6335 19	US-10-486-706-302 Sequence 302, App
44	237.2	3.8	5372 18	US-10-702-496-5 Sequence 5, Appl1
45	235.2	3.8	4917 18	US-10-385-163-121 Sequence 121, App

ALIGNMENTS

RESULT 1
US-10-028-946-1
Sequence 1, Application US/10028946
Publication No. US20020123622A1
GENERAL INFORMATION:
APPLICANT: Yu, Xuanchuan
APPLICANT: Miranda, Maricar
TITLE OF INVENTION: Fiddle, Carl Johan
FILE REFERENCE: LEX-0289-USA
CURRENT APPLICATION NUMBER: US/10/028,946
CURRENT FILING DATE: 2001-12-20
PRIOR APPLICATION NUMBER: US 60/256,335
PRIOR FILING DATE: 2000-12-27
NUMBER OF SEQ ID NOS: 4
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1
LENGTH: 6165
ORGANISM: homo sapiens
US-10-028-946-1

Query Match 100.0%; Score 6165; DB 13; Length 6165;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 6165; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGTTGAAGTTCAAAATATGAGCGCGGAATCTTGTGATGCTGCTGAACCAATT 60
Db 1 ATGTTGAAGTTCAAAATATGAGCGCGGAATCTTGTGATGCTGCTGAACCAATT 60
Qy 61 GCCAGCGGGGCTCCAGGCTGAATCTTCTTCAGGGAACCACTTATGACTCAA 120
Db 61 GCCAGCGGGGCTCCAGGCTGAATCTTCTTCAGGGAACCACTTATGACTCAA 120

D 2281 CAGCACTATGAGAAAAAGATTAAAGTGTGGACATCATGATTAAGAAAGACCTGGAC 2340
Q 2341 AAGAGACACTGAGAACATGATGACAGAGACAGAGAGAGGCCCATGAGAGGGGCAAA 2400
D 2341 AAGAGACACTGAGAACATGATGACAGAGACAGAGAGAGGCCCATGAGAGGGGCAAA 2400
Q 2401 ATTCTCAGCGAAGAGAGGGGATGATCATGCTATGGAATTCGAAATCAGATCCCTGGAA 2460
D 2401 ATTCTCAGCGAAGAGAGGGGATGATCATGCTATGGAATTCGAAATCAGATCCCTGGAA 2460
Q 2461 CAGAGATTGTGGAATGCTGTGAGAGCCAAATTAACCTTGACCAATAGAGCTTTTACC 2520
D 2461 CAGAGATTGTGGAATGCTGTGAGAGCCAAATTAACCTTGACCAATAGAGCTTTTACC 2520
Q 2521 CAAAGGAACATGAGAGGCCCAAGAGAGATGATTTCTGAATCTCAGGCAACAGAAATTTTAC 2580
D 2521 CAAAGGAACATGAGAGGCCCAAGAGAGATGATTTCTGAATCTCAGGCAACAGAAATTTTAC 2580
Q 2581 CTGGAGACACAGGCTGGAGAGTTGAGAGCCAGAAACGAAACCTGAGAGCAGCTGGAG 2640
D 2581 CTGGAGACACAGGCTGGAGAGTTGAGAGCCAGAAACGAAACCTGAGAGCAGCTGGAG 2640
Q 2641 AAGATCAGCCCAAGAGCCACAGTGAAGATGGGCTGGAACCTGGAACAAGATG 2700
D 2641 AAGATCAGCCCAAGAGCCACAGTGAAGATGGGCTGGAACCTGGAACAAGATG 2700
Q 2701 CGGAGAGCTCAGTCTAGAGACAGAGAGCAGAACTGAGCTCAAGCCGACGCTCAGAG 2760
D 2701 CGGAGAGCTCAGTCTAGAGACAGAGAGCAGAACTGAGCTCAAGCCGACGCTCAGAG 2760
Q 2761 CTACAGCTCTCCCTGAGAGAGCGGAGTCAAGTTGACAGCCCTGACAGGCTGCA 2820
D 2761 CTACAGCTCTCCCTGAGAGAGCGGAGTCAAGTTGACAGCCCTGACAGGCTGCA 2820
Q 2821 GCCCTGAGAGCCGAGCTTCCGACAGCCGAAAGACAGAGCTGGAAGAGACCAAGAGAGCT 2880
D 2821 GCCCTGAGAGCCGAGCTTCCGACAGCCGAAAGACAGAGCTGGAAGAGACCAAGAGAGCT 2880
Q 2881 GAAGAGAGATCCAGGCACTCAAGGCAATAGAGATGAATCCAGCCGAATTTGATGCT 2940
D 2881 GAAGAGAGATCCAGGCACTCAAGGCAATAGAGATGAATCCAGCCGAATTTGATGCT 2940
Q 2941 CTTGCTAACAGCTGTAATCTGTAATCAGACCTGAGAGAGCCTGAAACAGCTGACGAG 3000
D 2941 CTTGCTAACAGCTGTAATCTGTAATCAGACCTGAGAGAGCCTGAAACAGCTGACGAG 3000
Q 3001 GACAACTGTAATCTCAACCAAACTTCTTCTGTCGAAACCACTGATGAGGCTTCT 3060
D 3001 GACAACTGTAATCTCAACCAAACTTCTTCTGTCGAAACCACTGATGAGGCTTCT 3060
Q 3061 GGGGCAACGAGAGATTGTCAACTGAGAGTGAAGTGAACATCTCGGCGGGAGATC 3120
D 3061 GGGGCAACGAGAGATTGTCAACTGAGAGTGAAGTGAACATCTCGGCGGGAGATC 3120
Q 3121 ACGGAAGAGATGACGCTTACAGCCAGAGAAAGCAAGATGAGGCTCTGAAAGCAACG 3180
D 3121 ACGGAAGAGATGACGCTTACAGCCAGAGAAAGCAAGATGAGGCTCTGAAAGCAACG 3180
Q 3181 TGCACTGCTGAGAGAAAGAGTCAATGATTTGAGAGCCCTTAAACGATGAGCTGTGAA 3240
D 3181 TGCACTGCTGAGAGAAAGAGTCAATGATTTGAGAGCCCTTAAACGATGAGCTGTGAA 3240
Q 3241 AAAAGAGGAGAGTGGAGAGGCTGAGAGAGCTGAGGATGAGAAATCCCAAGTTTAC 3300
D 3241 AAAAGAGGAGAGTGGAGAGGCTGAGAGAGCTGAGGATGAGAAATCCCAAGTTTAC 3300
Q 3301 TGTGAGGTTGAGAGCTGAGAGAGTGTGAGCAACGAGAAACAGAGCGAGGCGAGAGCC 3360
D 3301 TGTGAGGTTGAGAGCTGAGAGAGTGTGAGCAACGAGAAACAGAGCGAGGCGAGAGCC 3360
Q 3361 GATCAGCGGATCCAGAGCTTCCGAGGCTGAGAGCTGAGAGAGCAAGAGCT 3420
D 3361 GATCAGCGGATCCAGAGCTTCCGAGGCTGAGAGCTGAGAGAGCAAGAGCT 3420

Q 3421 GAGATTCTCGCTGACAGCGCTCTCAAGAGAGAGAGCTGAAGGCCGAGAGCTCTCT 3480
D 3421 GAGATTCTCGCTGACAGCGCTCTCAAGAGAGAGAGCTGAAGGCCGAGAGCTCTCT 3480
Q 3481 GACAGCTCAATGACCTGAGAGAGAGCAATGCTTGTGTAATGAAATGAGCCGAACTTA 3540
D 3481 GACAGCTCAATGACCTGAGAGAGAGCAATGCTTGTGTAATGAAATGAGCCGAACTTA 3540
Q 3541 CAGCAGAGCTGAGAGCTGAAAGAGAGCTCAACAGAGGCTTCTGAAAGAGCAAGCCAA 3600
D 3541 CAGCAGAGCTGAGAGCTGAAAGAGAGCTCAACAGAGGCTTCTGAAAGAGCAAGCCAA 3600
Q 3601 TTAAGAGAGATGAGCTGAGAGAAATACATTTTCCGCTGACCTCAAGAGCTGCA 3660
D 3601 TTAAGAGAGATGAGCTGAGAGAAATACATTTTCCGCTGACCTCAAGAGCTGCA 3660
Q 3661 GAAGCTTGAATGGGCTGATCTATCTGAAGACAGAAAGAGTGAAGCTTGAAGATCAGCTG 3720
D 3661 GAAGCTTGAATGGGCTGATCTATCTGAAGACAGAAAGAGTGAAGCTTGAAGATCAGCTG 3720
Q 3721 GAAGACATTCAGGTTCTCTATTTCTCATGAAAGGTTGAAGAGCACTTATTTCTCA 3780
D 3721 GAAGACATTCAGGTTCTCTATTTCTCATGAAAGGTTGAAGAGCACTTATTTCTCA 3780
Q 3781 CAAACCAACTCATTTATTTTCTGCAAGCCAAATGAGCAACCTGCTAAAAAGAAAG 3840
D 3781 CAAACCAACTCATTTATTTTCTGCAAGCCAAATGAGCAACCTGCTAAAAAGAAAG 3840
Q 3841 GTTCTCTGACATCATGAGCTGAAAGCTGGCCCTGAGAGAGAGAAAGCTCGCTGCA 3900
D 3841 GTTCTCTGACATCATGAGCTGAAAGCTGGCCCTGAGAGAGAGAAAGCTCGCTGCA 3900
Q 3901 GAGCTAGAGAGAGGCTTCTGAGAGAGCCGATCGAGCTCGGCTCGCCGGAGAGAGCT 3960
D 3901 GAGCTAGAGAGAGGCTTCTGAGAGAGCCGATCGAGCTCGGCTCGCCGGAGAGAGCT 3960
Q 3961 GAGCTAGAGAGAGGCTTCTGAGAGAGCCGATCGAGCTCGGCTCGCCGGAGAGAGCT 3960
D 3961 GAGCTAGAGAGAGGCTTCTGAGAGAGCCGATCGAGCTCGGCTCGCCGGAGAGAGCT 3960
Q 4001 GAGCTAGAGAGAGGCTTCTGAGAGAGCCGATCGAGCTCGGCTCGCCGGAGAGAGCT 3960
D 4001 GAGCTAGAGAGAGGCTTCTGAGAGAGCCGATCGAGCTCGGCTCGCCGGAGAGAGCT 3960
Q 4021 ATCCCATGCTCGGATCGGCTGAGAGAGCCGATCGAGCTCGGCTCGCCGGAGAGAGCT 4080
D 4021 ATCCCATGCTCGGATCGGCTGAGAGAGCCGATCGAGCTCGGCTCGCCGGAGAGAGCT 4080
Q 4081 GAGCTAGAGAGAGGCTTCTGAGAGAGCCGATCGAGCTCGGCTCGCCGGAGAGAGCT 4140
D 4081 GAGCTAGAGAGAGGCTTCTGAGAGAGCCGATCGAGCTCGGCTCGCCGGAGAGAGCT 4140
Q 4141 AAGGAGAGAGAGGCTTCTGAGAGAGCCGATCGAGCTCGGCTCGCCGGAGAGAGCT 4200
D 4141 AAGGAGAGAGAGGCTTCTGAGAGAGCCGATCGAGCTCGGCTCGCCGGAGAGAGCT 4200
Q 4201 ACAAAGTGTGCTGTGTCTGAGATCGGCTGAGAGAGCCGATCGAGCTCGGCTCGCC 4260
D 4201 ACAAAGTGTGCTGTGTCTGAGATCGGCTGAGAGAGCCGATCGAGCTCGGCTCGCC 4260
Q 4261 GAAATGATGAGAGAGGCTTCAACCAAGTGTCAAGGCTTCAAGGCTTCAAGGCTTCA 4320
D 4261 GAAATGATGAGAGAGGCTTCAACCAAGTGTCAAGGCTTCAAGGCTTCAAGGCTTCA 4320
Q 4321 CTTGCTGAATATGACACACTTCAACGAGGCTTCTGAGAGAGAGCTTCAAGGCTTCA 4380
D 4321 CTTGCTGAATATGACACACTTCAACGAGGCTTCTGAGAGAGAGCTTCAAGGCTTCA 4380
Q 4381 GGTCTCAGAGAGAGAGGCTTCAAGAGAGCTTCAAGAGAGAGGCTTCAAGAGAGGCT 4440
D 4381 GGTCTCAGAGAGAGAGGCTTCAAGAGAGCTTCAAGAGAGAGGCTTCAAGAGAGGCT 4440
Q 4441 AGGAAATTAACAAAGAGAGAGGCTTCAAGAGAGGCTTCAAGAGAGAGGCTTCAAGAG 4500
D 4441 AGGAAATTAACAAAGAGAGAGGCTTCAAGAGAGGCTTCAAGAGAGAGGCTTCAAGAG 4500

4501 AAAGTCTCATTTATGACATGAAAGCCAGAGAGCTGAGCAGAGGCGGTGGAAGATT 4560
4501 AAAGTCTCATTTATGACATGAAAGCCAGAGAGCTGAGCAGAGGCGGTGGAAGATT 4560
4561 GAGCTGTGCTTCCGAGCGGGATGTATCTATTATGAGTGCCTGTGTGCTTCCGAATC 4620
4561 GAGCTGTGCTTCCGAGCGGGATGTATCTATTATGAGTGCCTGTGTGCTTCCGAATC 4620
4621 GCAAAATACGCAAGAGAGATGTCCTCATCATCTGAAGATGAAATCTCACCCGACAC 4680
4621 GCAAAATACGCAAGAGAGATGTCCTCATCATCTGAAGATGAAATCTCACCCGACAC 4680
4681 ACCTGTGCGCCGGAGAGACCTCTACTGTAGTCCAGCTTCCCGAACAACAGGCG 4740
4681 ACCTGTGCGCCGGAGAGACCTCTACTGTAGTCCAGCTTCCCGAACAACAGGCG 4740
4741 TGGGTCAACCGCTTAGATCATGTGTGAGTGGAGTGGAGATTTCTAGGGAAGAGAGAA 4800
4741 TGGGTCAACCGCTTAGATCATGTGTGAGTGGAGTGGAGATTTCTAGGGAAGAGAGAA 4800
4801 GCTGATGCTAACTGCTTTGAGAACTCCCTGTGAACTGGAGGTGATGACCTGTAGAC 4860
4801 GCTGATGCTAACTGCTTTGAGAACTCCCTGTGAACTGGAGGTGATGACCTGTAGAC 4860
4861 ATGAACTGACAGCTGCTTCAAGTACAGTGTGAGTGGAGTGGAGTGGAGTGGAGTGG 4920
4861 ATGAACTGACAGCTGCTTCAAGTACAGTGTGAGTGGAGTGGAGTGGAGTGGAGTGG 4920
4921 TACGCTGATGATGCTTTGAAAGATCCCTTAACCCATGCTCCAGAAATGAGAGCTTTC 4980
4921 TACGCTGATGATGCTTTGAAAGATCCCTTAACCCATGCTCCAGAAATGAGAGCTTTC 4980
4981 GAAATTTATTTATCAAGAGCTGAGAGAGTACTCATGATAGAGAGAGAGAGAGAGAG 5040
4981 GAAATTTATTTATCAAGAGCTGAGAGAGTACTCATGATAGAGAGAGAGAGAGAGAG 5040
5041 CTGTGCTTGTGAGAGTGAAGAAAGTGAACAGTCCCTGAGCCAGTCCCACTGCTGCC 5100
5041 CTGTGCTTGTGAGAGTGAAGAAAGTGAACAGTCCCTGAGCCAGTCCCACTGCTGCC 5100
5101 CAGCCGACACTCTACCCCAATTTTGAAGCTGTCAAGGCTGCCACTGTTTGGGCA 5160
5101 CAGCCGACACTCTACCCCAATTTTGAAGCTGTCAAGGCTGCCACTGTTTGGGCA 5160
5161 GCGAAGATTGAGAACGGGCTCTGCACTGTGAGCCATGCTCCAGCAAAAGTCTTCTC 5220
5161 GCGAAGATTGAGAACGGGCTCTGCACTGTGAGCCATGCTCCAGCAAAAGTCTTCTC 5220
5221 CGCTACAGGAAACCTCAGCAAAATCTGATCCGAGAAAGATAGAGACTCAGAGCCC 5280
5221 CGCTACAGGAAACCTCAGCAAAATCTGATCCGAGAAAGATAGAGACTCAGAGCCC 5280
5281 TGCAGCTGATCACTTCAACCAATTCAGTATCTCATTTGAGAACCAATTAATCTACGA 5340
5281 TGCAGCTGATCACTTCAACCAATTCAGTATCTCATTTGAGAACCAATTAATCTACGA 5340
5341 ATGACATGAGAGAGTACAGCTCGAGGAATTCCTGATTAAGATGACCAATTCCTTGGCA 5400
5341 ATGACATGAGAGAGTACAGCTCGAGGAATTCCTGATTAAGATGACCAATTCCTTGGCA 5400
5401 CCGTGTGTTTGGCGGCTCTTCCACAGCTTCCCTGTCTCAATCTGTGAGAGTGAACAGC 5460
5401 CCGTGTGTTTGGCGGCTCTTCCACAGCTTCCCTGTCTCAATCTGTGAGAGTGAACAGC 5460
5461 GCGAGGACGAGAGAGTACTGTGTGTTTCCAGAAATTTGAGAGTGTGAGATTC 5520
5461 GCGAGGACGAGAGAGTACTGTGTGTTTCCAGAAATTTGAGAGTGTGAGATTC 5520
5521 TACGGAAGCTGAGCCGACAGACGATCTCAAGTGAAGTGGCTTACCTTTGGCTTTGCC 5580
5521 TACGGAAGCTGAGCCGACAGACGATCTCAAGTGAAGTGGCTTACCTTTGGCTTTGCC 5580
5581 TACAGAGAACCTATCTGTTTGTGAGCCCACTTCACTCACTGGAAGTAAATGAGATCCAG 5640

5581 TACAGAGAACCTATCTGTTTGTGAGCCCACTTCACTCACTGGAAGTAAATGAGATCCAG 5640
5641 GCAGCTCTCTCAGAGAGAGACCTCTGCGGAGCTGAGACATCTCGAATCCCGCTAC 5700
5641 GCAGCTCTCTCAGAGAGAGACCTCTGCGGAGCTGAGACATCTCGAATCCCGCTAC 5700
5701 CTGGGCTCTGCTTCTTCTCAGAGAGATTTACTTGGCTGCTCTCATACCAAGATTAATTA 5760
5701 CTGGGCTCTGCTTCTTCTCAGAGAGATTTACTTGGCTGCTCTCATACCAAGATTAATTA 5760
5761 AGGCTCATTTGTGCAAGAGAGAACTGTGAAGAGTCCGCACTGAACCAACCGGCGC 5820
5761 AGGCTCATTTGTGCAAGAGAGAACTGTGAAGAGTCCGCACTGAACCAACCGGCGC 5820
5821 CCGTCACTCTCCGACAGACCCCAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5880
5821 CCGTCACTCTCCGACAGACCCCAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5880
5881 ACCAAGCGCTGAGCTTCCAGCCAGCGCGCTGGAAGAGAGAGAGAGAGAGAGAGAG 5940
5881 ACCAAGCGCTGAGCTTCCAGCCAGCGCGCTGGAAGAGAGAGAGAGAGAGAGAGAG 5940
5941 AGGACACCCCAACCGCTACCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6000
5941 AGGACACCCCAACCGCTACCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6000
6001 CCGCCCTGAG 6060
6001 CCGCCCTGAG 6060
6061 CCGCGAG 6120
6061 CCGCGAG 6120
6121 CCGCTGTCCAGAGTGAACAGGCTGTGAGACAGATCTTCACTATTA 6165
6121 CCGCTGTCCAGAGTGAACAGGCTGTGAGACAGATCTTCACTATTA 6165

RESULT 2
US-10-791-666-1
Sequence 1, Application US/10791666
Publication No. US20040209297A1
GENERAL INFORMATION:
APPLICANT: Yu, Xuanhuan
APPLICANT: Miranda, Maricar
APPLICANT: Fiedler, Carl Johan
TITLE OF INVENTION: Novel Human Kinases and Polynucleotides Encoding the Same
FILE REFERENCE: LEX-0289-USA
CURRENT APPLICATION NUMBER: US/10/791,666
PRIOR FILING DATE: 2004-03-02
PRIOR APPLICATION NUMBER: US/10/028,946
PRIOR FILING DATE: 2001-12-20
PRIOR APPLICATION NUMBER: US 60/258,335
NUMBER OF SEQ ID NOS: 4
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1
LENGTH: 6165
TYPE: DNA
ORGANISM: homo sapiens
US-10-791-666-1

Query Match 100.0%; Score 6165; DB 18; Length 6165;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 6165; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 ATGTTGAAGTTCAATATGAGAGCGGGAATCTTTGAGATGCTGTGCTGTAACCCATT 60
1 ATGTTGAAGTTCAATATGAGAGCGGGAATCTTTGAGATGCTGTGCTGTAACCCATT 60
61 GCGAGCCGGGCTCCAGGCTGATCTGTTTCCAGGAGAGAGAGAGAGAGAGAGAGAGAG 120

Db 61 GCACGCGGGCCTCCAGCTGAACTGTTCTTCCAGGGGAAACACCCCTTATGACTCAA 120
Qy 121 CAGAGATGTCCTCTCTTCCCGAGAGGATATTAGATGCCCTCTTGTGTTCTTTGGAA 180
Db 121 CAGAGATGTCCTCTCTTCCCGAGAGGATATTAGATGCCCTCTTGTGTTCTTTGGAA 180
Qy 181 GAATGCAGTCACTGCTGTGATGAGATTAAAGCAGTGACAACTTGTCCGAGATAT 240
Db 181 GAATGCAGTCACTGCTGTGATGAGATTAAAGCAGTGACAACTTGTCCGAGATAT 240
Qy 241 TCCGACACCATAGCTGATTACAGAGAGCTCCAGCTTCCGCAAGAGCTTGAAGTCA 300
Db 241 TCCGACACCATAGCTGATTACAGAGAGCTCCAGCTTCCGCAAGAGCTTGAAGTCA 300
Qy 301 AGCTTTAGGTTGTGCTGCTTGTGGAAGTGAGGCTGTAAGAGAAAGCAACCGGG 360
Db 301 AGCTTTAGGTTGTGCTGCTTGTGGAAGTGAGGCTGTAAGAGAAAGCAACCGGG 360
Qy 361 GACATCTATGCTATGAAAGTATGAGAGAGAGCTTATTTGCCCAGAGCAGTTTCA 420
Db 361 GACATCTATGCTATGAAAGTATGAGAGAGAGCTTATTTGCCCAGAGCAGTTTCA 420
Qy 421 TTTTGAAGAGAGCGGAAATATTAATCTGAGACAAGCCCGTGATCCCAATTA 480
Db 421 TTTTGAAGAGAGCGGAAATATTAATCTGAGACAAGCCCGTGATCCCAATTA 480
Qy 481 CAGATGCTTTGAGCAAAATATCCTTATCTGTGCAATGAAATATGAGCTGGAGGG 540
Db 481 CAGATGCTTTGAGCAAAATATCCTTATCTGTGCAATGAAATATGAGCTGGAGGG 540
Qy 541 GACTGCTGTCATTTTGAATAGATATGAGAGCAGTTAGTGAATCACTGATCACTTT 600
Db 541 GACTGCTGTCATTTTGAATAGATATGAGAGCAGTTAGTGAATCACTGATCACTTT 600
Qy 601 TACCTAGCTGAGCTGATTTTGGCTGTTCAAGCGCTTCACTGATGGAGATCGTCATCGA 660
Db 601 TACCTAGCTGAGCTGATTTTGGCTGTTCAAGCGCTTCACTGATGGAGATCGTCATCGA 660
Qy 661 GAGATCAAGCTGAGCAATCTGTTGACCGGACAGGACATCAAGCTGGTGGATTTT 720
Db 661 GAGATCAAGCTGAGCAATCTGTTGACCGGACAGGACATCAAGCTGGTGGATTTT 720
Qy 721 GATCTGCTGCGGAAATGAAATTCAAACAGATGCTGAATGCCAATCTCCGATTTGGAGC 780
Db 721 GATCTGCTGCGGAAATGAAATTCAAACAGATGCTGAATGCCAATCTCCGATTTGGAGC 780
Qy 781 CCAAGATTACATGCTGCTGAAAGTCTGAATGATGAAAGGGGATGAAAGGCACTTAC 840
Db 781 CCAAGATTACATGCTGCTGAAAGTCTGAATGATGAAAGGGGATGAAAGGCACTTAC 840
Qy 841 GGGCTGGAATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 900
Db 841 GGGCTGGAATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 900
Qy 901 TCCCTCTGCGAGAGGGAACCTGCGCAGAACCTTCAATTAATTAATTAATTTCCAGGG 960
Db 901 TCCCTCTGCGAGAGGGAACCTGCGCAGAACCTTCAATTAATTAATTAATTTCCAGGG 960
Qy 961 TTTTGAATTTCCAGATGACCCCAAGTGAGCAGTACTTTCTTGATCTGATTCAAAGC 1020
Db 961 TTTTGAATTTCCAGATGACCCCAAGTGAGCAGTACTTTCTTGATCTGATTCAAAGC 1020
Qy 1021 TTGTTGTCGCGGCGAGAAAGAGAGCTGAAGTTTGAAGTCTTTGCTGCACTCTTCTTC 1080
Db 1021 TTGTTGTCGCGGCGAGAAAGAGAGCTGAAGTTTGAAGTCTTTGCTGCACTCTTCTTC 1080
Qy 1081 TCTAAATTTGATGAGAAACAATTCGTAATCTCTCCCTCTGTTCCCACTCTCAAG 1140
Db 1081 TCTAAATTTGATGAGAAACAATTCGTAATCTCTCCCTCTGTTCCCACTCTCTCAAG 1140
Qy 1141 TCTGACATGACACTTCAATTTTGAATGACAGAGAAATTCGTGGGTTTCACTCTCT 1200
Db 1141 TCTGACATGACACTTCAATTTTGAATGACAGAGAAATTCGTGGGTTTCACTCTCT 1200

Db 1141 TCTGACATGACACTTCAATTTTGAATGACAGAGAAATTCGTGGGTTTCACTCTCT 1200
Qy 1201 CCGTGCCAGCTGAGCCCTCAGGCTTCTGGGTGAAGAACTGCGTTTGTGGGGTTTTCG 1260
Db 1201 CCGTGCCAGCTGAGCCCTCAGGCTTCTGGGTGAAGAACTGCGTTTGTGGGGTTTTCG 1260
Qy 1261 TACAGCAAGGCACTGGGGATTTCTTGTGATCTGAAGTCTGTGTGTGGGTCTGCACTCC 1320
Db 1261 TACAGCAAGGCACTGGGGATTTCTTGTGATCTGAAGTCTGTGTGTGGGTCTGCACTCC 1320
Qy 1321 CTTGCGCAAGCTAGCTGAGAAAGAACTTCTCATCAAAAGCAAGAGCTCAAGAC 1380
Db 1321 CTTGCGCAAGCTAGCTGAGAAAGAACTTCTCATCAAAAGCAAGAGCTCAAGAC 1380
Qy 1381 TCTCAGAGCAAGTGTCAAGATGAGCAGAGAAATGAACCGGTTTACATCGGAGTGTCA 1440
Db 1381 TCTCAGAGCAAGTGTCAAGATGAGCAGAGAAATGAACCGGTTTACATCGGAGTGTCA 1440
Qy 1441 GAGGTGAGGCTGTGCTTAACTGAGAGAGAGTGAAGGCTTGAAGCTTCAAGAGA 1500
Db 1441 GAGGTGAGGCTGTGCTTAACTGAGAGAGAGTGAAGGCTTGAAGCTTCAAGAGA 1500
Qy 1501 TCCCTCTGAGAGAGAGCTTGTGCTAATCAAGATGCAAGATGCAAGCTTAAAGCAGAGT 1560
Db 1501 TCCCTCTGAGAGAGAGCTTGTGCTAATCAAGATGCAAGATGCAAGCTTAAAGCAGAGT 1560
Qy 1561 TTGAGAGCAAGCAGATGAGAGTGTCCAGAGAGATGACAAAGCACTGACGCTTCTCAT 1620
Db 1561 TTGAGAGCAAGCAGATGAGAGTGTCCAGAGAGATGACAAAGCACTGACGCTTCTCAT 1620
Qy 1621 GATATCAGAGAGCAGAGCCGGAAGCTTCAAGAAATCAAGAGCAGAGATACAGGCTCAA 1680
Db 1621 GATATCAGAGAGCAGAGCCGGAAGCTTCAAGAAATCAAGAGCAGAGATACAGGCTCAA 1680
Qy 1681 GTGGAAGAAATGAGGTTGATGATGAATCAAGTTGGAAGAGATCTTGTCTGCAAGAGAA 1740
Db 1681 GTGGAAGAAATGAGGTTGATGATGAATCAAGTTGGAAGAGATCTTGTCTGCAAGAGAA 1740
Qy 1741 CCGAGTATCTTACGATCTGAGCTGAGAGAGTCTGGGCTTGGCTGAGAAATTCAG 1800
Db 1741 CCGAGTATCTTACGATCTGAGCTGAGAGAGTCTGGGCTTGGCTGAGAAATTCAG 1800
Qy 1801 CCGAAGCAGACAGATGTCAAGATTAATCTGTTAAGGCTTAAGATCAAGGAGCTTGA 1860
Db 1801 CCGAAGCAGACAGATGTCAAGATTAATCTGTTAAGGCTTAAGATCAAGGAGCTTGA 1860
Qy 1861 GTGGAGAAATGTGGAAATCTGAGAGAGATCAATCTGAGCAGAGCTTCAAAATTCAGAG 1920
Db 1861 GTGGAGAAATGTGGAAATCTGAGAGAGATCAATCTGAGCAGAGCTTCAAAATTCAGAG 1920
Qy 1921 CTCGAGAGAAATCTGAGAGAGCTGTAAAGCTGTAAGCTGAGAGCTGAGAGCTGAGAG 1980
Db 1921 CTCGAGAGAAATCTGAGAGAGCTGTAAAGCTGTAAGCTGAGAGCTGAGAGCTGAGAG 1980
Qy 1981 AATATCGGCAAGGCAAG 2040
Db 1981 AATATCGGCAAGGCAAG 2040
Qy 2041 GATTTCTTGAAGCATTCAAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2100
Db 2041 GATTTCTTGAAGCATTCAAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2100
Qy 2101 GAGAAAGAGGTAAG 2160
Db 2101 GAGAAAGAGGTAAG 2160
Qy 2161 ATCCAGACAAATTCAG 2220
Db 2161 ATCCAGACAAATTCAG 2220
Qy 2221 AAACATGGGAGGCTCAAGTCTCAGCCAGCACTAGAGAGAGAGAGAGAGAGAGAGAG 2280
Db 2221 AAACATGGGAGGCTCAAGTCTCAGCCAGCACTAGAGAGAGAGAGAGAGAGAGAGAG 2280

QY 2281 CAGCACTATGAGAAAGATTAAAGTGTGGCAATCATCATTAAGAAAGACCTGGCTGAC 2340
DB 2281 CAGCACTATGAGAAAGATTAAAGTGTGGCAATCATCATTAAGAAAGACCTGGCTGAC 2340
QY 2341 AAGGAGACACTGAGAGAAATGATGACAGACACGAGAGAGAGGCCCATGAGAGAGGAGAA 2400
DB 2341 AAGGAGACACTGAGAGAAATGATGACAGACACGAGAGAGAGGCCCATGAGAGAGGAGAA 2400
QY 2401 ATTCTACCGCAACGAAGGCGATGATCAATGCTATGGAATTCGAAGATCAGATCCCTGAA 2460
DB 2401 ATTCTACCGCAACGAAGGCGATGATCAATGCTATGGAATTCGAAGATCAGATCCCTGAA 2460
QY 2461 CAGAGATTGTGGAATCTGTCTGAAGCCAAATTAACCTTGACGAATATGACGCTTTTAAAC 2520
DB 2461 CAGAGATTGTGGAATCTGTCTGAAGCCAAATTAACCTTGACGAATATGACGCTTTTAAAC 2520
QY 2521 CAAAGGAACATGAAGGCCCAAGAGATGATTTCTGAATCTAGGCAACAGAAATTTTAC 2580
DB 2521 CAAAGGAACATGAAGGCCCAAGAGATGATTTCTGAATCTAGGCAACAGAAATTTTAC 2580
QY 2581 CTGGAGACACAGGCTGGAAAGTTGAGAGCCCAAGAACGAAACTGAGAGACAGCTGAG 2640
DB 2581 CTGGAGACACAGGCTGGAAAGTTGAGAGCCCAAGAACGAAACTGAGAGACAGCTGAG 2640
QY 2641 AAGATCGCAACCAAGACCAAGTGAACAAGATTCGCTGTGGAATCTGAGACAAAGTTG 2700
DB 2641 AAGATCGCAACCAAGACCAAGTGAACAAGATTCGCTGTGGAATCTGAGACAAAGTTG 2700
QY 2701 CCGAGAGTCACTGATGAGACAGAGACAGAACTGAGAGCTCAAGGCGCTCAAGAG 2760
DB 2701 CCGAGAGTCACTGATGAGACAGAGACAGAACTGAGAGCTCAAGGCGCTCAAGAG 2760
QY 2761 CTACAGCTCTCCCTGAGAGCGGAGTCAAGTTGACAGCCCTGACGCTGACCGGCG 2820
DB 2761 CTACAGCTCTCCCTGAGAGCGGAGTCAAGTTGACAGCCCTGACGCTGACCGGCG 2820
QY 2821 GCCCTGAGAGCCAGCTTCCGCGAGGCAAGACAGAGCTGGAAGAGACCAAGAAAGCT 2880
DB 2821 GCCCTGAGAGCCAGCTTCCGCGAGGCAAGACAGAGCTGGAAGAGACCAAGAAAGCT 2880
QY 2881 GAAGAGAGATCCAGGCACTCAGGCGACATGAGATGGAATCCAGGCGCAATTTGATGCT 2940
DB 2881 GAAGAGAGATCCAGGCACTCAGGCGACATGAGATGGAATCCAGGCGCAATTTGATGCT 2940
QY 2941 CTTCTGTAACAGCTGTA CTGTATCAAGACCTGAGAGAGAGCTTAAACAGCTGACGAG 3000
DB 2941 CTTCTGTAACAGCTGTA CTGTATCAAGACCTGAGAGAGAGCTTAAACAGCTGACGAG 3000
QY 3001 GACACGCTGTAATCAACCAAACTTTACTTGTCCAAACAATCTCGATGAGGCTTCT 3060
DB 3001 GACACGCTGTAATCAACCAAACTTTACTTGTCCAAACAATCTCGATGAGGCTTCT 3060
QY 3061 GCGCGCAACGAGAGATTGTAACACTGCGAAGTGAAGTGAACAATCTCGCGCGGAGATC 3120
DB 3061 GCGCGCAACGAGAGATTGTAACACTGCGAAGTGAAGTGAACAATCTCGCGCGGAGATC 3120
QY 3121 ACGGAAACGAGAGTGAAGCTTACACGCAAGAACAAAGATGAGGCTGTGAAGACCAAG 3180
DB 3121 ACGGAAACGAGAGTGAAGCTTACACGCAAGAACAAAGATGAGGCTGTGAAGACCAAG 3180
QY 3181 TGCACATCTGAGAGAAACAGTCAATGATTTGAGAGCCCTTAAACGATGAGCTGTAGAA 3240
DB 3181 TGCACATCTGAGAGAAACAGTCAATGATTTGAGAGCCCTTAAACGATGAGCTGTAGAA 3240
QY 3241 AAAGAGCGCAGTGGAGGCTTGAAGAGAGGCTCTGGGTGATGAAGAAATCCAGTTTGAAG 3300
DB 3241 AAAGAGCGCAGTGGAGGCTTGAAGAGAGGCTCTGGGTGATGAAGAAATCCAGTTTGAAG 3300
QY 3301 TGTGCGGTTTGAAGCTGAGAGATGTCTGACACCGAGAAACAGAGCAGGCGGAGAGGCC 3360
DB 3301 TGTGCGGTTTGAAGCTGAGAGATGTCTGACACCGAGAAACAGAGCAGGCGGAGAGGCC 3360

QY 3361 GATCAGCGGATCACCGAGTCTCCGACAGGTGTGAGCTGCGCATGAAGAGACCAAGGCT 3420
DB 3361 GATCAGCGGATCACCGAGTCTCCGACAGGTGTGAGCTGCGCATGAAGAGACCAAGGCT 3420
QY 3421 GAGATTCTGCTGCGAGCGGCTCTCAAGAGAGAGAGCTGAAGGCGGAGAGGCTCTCT 3480
DB 3421 GAGATTCTGCTGCGAGCGGCTCTCAAGAGAGAGAGCTGAAGGCGGAGAGGCTCTCT 3480
QY 3481 GACAACTCAATGACTTGGAGAGAGACATGCTTATGCTTGAATGAATGATGCTCGAAGCTTA 3540
DB 3481 GACAACTCAATGACTTGGAGAGAGACATGCTTATGCTTGAATGAATGATGCTCGAAGCTTA 3540
QY 3541 CAGCAGAACTGGAAGCTGAACGAGAGCTCAACAGAGGCTTCTGGAAGAGCAAGCCAA 3600
DB 3541 CAGCAGAACTGGAAGAGTGAACGAGAGCTCAACAGAGGCTTCTGGAAGAGCAAGCCAA 3600
QY 3601 TTACAGCAGAGATGAGCTGCGAAGAAATCACTTTCGCTGACTCAAGGACTGCA 3660
DB 3601 TTACAGCAGAGATGAGCTGCGAAGAAATCACTTTCGCTGACTCAAGGACTGCA 3660
QY 3661 GAAGCTCTAGATCGGCTGATCTTATGAGACAGAAAGAAAGTGAATCTGAGATACAGCTG 3720
DB 3661 GAAGCTCTAGATCGGCTGATCTTATGAGACAGAAAGAAAGTGAATCTGAGATACAGCTG 3720
QY 3721 GAAACATTGAGGTTCTTATTTCTCATGAAGAGTGAATGAAGGCACTATTTCTCAA 3780
DB 3721 GAAACATTGAGGTTCTTATTTCTCATGAAGAGTGAATGAAGGCACTATTTCTCAA 3780
QY 3781 CAAACCAACTCAATGATTTTCTGCAAGCCAAATGGAACCACTGCTGTAAGAAAG 3840
DB 3781 CAAACCAACTCAATGATTTTCTGCAAGCCAAATGGAACCACTGCTGTAAGAAAG 3840
QY 3841 GTTCTCTGAGTACATGAGTGAAGCTGAGCCCTGAGAGAGAGAAAGCTGCTGTGCA 3900
DB 3841 GTTCTCTGAGTACATGAGTGAAGCTGAGCCCTGAGAGAGAGAAAGCTGCTGTGCA 3900
QY 3901 GAGCTGAGAGAGCCCTTGAAGAGACCCGATGAGCTCGCTCCGCGGAGAGAGCT 3960
DB 3901 GAGCTGAGAGAGCCCTTGAAGAGACCCGATGAGCTCGCTCCGCGGAGAGAGCT 3960
QY 3961 GCCCACGGAAGCAACGAGACCAACCACTCCAGCCGACCCGCGAGAGGACGAG 4020
DB 3961 GCCCACGGAAGCAACGAGACCAACCACTCCAGCCGACCCGCGAGAGGACGAG 4020
QY 4021 ATGCGCATGTCCGCAATCGTGCAGTGCAGAGACCAAGCCAGTGCATGAGGCTGCTG 4080
DB 4021 ATGCGCATGTCCGCAATCGTGCAGTGCAGAGACCAAGCCAGTGCATGAGGCTGCTG 4080
QY 4081 GCCCGCAATCCAGCCGCAAGAAAGAGTCTTCAACTCCAGAGAAATTTAGTGGGCTT 4140
DB 4081 GCCCGCAATCCAGCCGCAAGAAAGAGTCTTCAACTCCAGAGAAATTTAGTGGGCTT 4140
QY 4141 AAGGAACGATGACACCAATTTCTTCTCAACGATTCAGATGAGACTGAACATGCGAGGC 4200
DB 4141 AAGGAACGATGACACCAATTTCTTCTCAACGATTCAGATGAGACTGAACATGCGAGGC 4200
QY 4201 ACAAAGTGTGCTGTGTGTGATACCGTGCATTGTGAAGCGCAGGCAATCAATGTCTC 4260
DB 4201 ACAAAGTGTGCTGTGTGTGATACCGTGCATTGTGAAGCGCAGGCAATCAATGTCTC 4260
QY 4261 GAATGTCAAGTGTGTGTCAACCCCAAGTGTCTCAAGTGTCTCAAGGCTGCGGCTTG 4320
DB 4261 GAATGTCAAGTGTGTGTGTCAACCCCAAGTGTGTCTCAAGTGTCTCAAGGCTGCGGCTTG 4320
QY 4321 CCTGCTGAATATGCAACACTTCAACGAGGCTTCTGCGGTGACAAATGAATCTCCCA 4380
DB 4321 CCTGCTGAATATGCAACACTTCAACGAGGCTTCTGCGGTGACAAATGAATCTCCCA 4380
QY 4381 GGTCTTCAGACCAAGAGCCGAGAGCTTGAAGCTTGAAGAGGTTGAAGAGGCTGCC 4440
DB 4381 GGTCTTCAGACCAAGAGCCGAGAGCTTGAAGCTTGAAGAGGTTGAAGAGGCTGCC 4440
QY 4441 AGGAATTAACAAACGAGACAGCAAGGCTGGGACAGAAAGTACATTTGTCTGAGAGGATCA 4500

```

Db      4441 AGGAMTAAACAAAGAGACAGCAGGCTGGACAGGAAGTACATTTGCTGGAGGAGTCA      4500
Qy      4501 AAAGTCTCATTTATGACAAATGAAAGCAGAGAACTGAGACAGGCGGTGAGAAATTT      4560
Db      4501 AAAGTCTCATTTATGACAAATGAAAGCAGAGAACTGAGACAGGCGGTGAGAAATTT      4560
Qy      4561 GAGTGTGCTTCCGACGCGGAGATGATCTATTGATGTCGCTTGGTCTTCCGAACTC      4620
Db      4561 GAGTGTGCTTCCGACGCGGAGATGATCTATTGATGTCGCTTGGTCTTCCGAACTC      4620
Qy      4621 GCAAATACAGCCAAAGCAGATGTCTCATCTATGAAAGATGAAATCTCACCCGACACC      4680
Db      4621 GCAAATACAGCCAAAGCAGATGTCTCATCTATGAAAGATGAAATCTCACCCGACACC      4680
Qy      4681 ACCGTGCGCCGCGGAGAACCTCTACTGCTAGCTCCAGCTTCCCTGACAAACAGGCG      4740
Db      4681 ACCGTGCGCCGCGGAGAACCTCTACTGCTAGCTCCAGCTTCCCTGACAAACAGGCG      4740
Qy      4741 TGGGTCAACGCTTGAATCAGTTGTGCAAGTGGAGAGTTTCTAGGGAAGAAAGCAGAA      4800
Db      4741 TGGGTCAACGCTTGAATCAGTTGTGCAAGTGGAGAGTTTCTAGGGAAGAAAGCAGAA      4800
Qy      4801 GCTGATCTTAACTGCTTGGAAATCTCCCTGCTGAACTGGAAGGTATGACCTGTAGAC      4860
Db      4801 GCTGATCTTAACTGCTTGGAAATCTCCCTGCTGAACTGGAAGGTATGACCTGTAGAC      4860
Qy      4861 ATGAACTGCAACGCTGCTTCAAGTACAGAGTGGTGGTGGGACCCGAGAGAGGCTC      4920
Db      4861 ATGAACTGCAACGCTGCTTCAAGTACAGAGTGGTGGTGGGACCCGAGAGAGGCTC      4920
Qy      4921 TAGCCCTGATGTCTTGAAGAACTCCCTAACCCATGCTCCAGAAATTTGAGCAGTCTTC      4980
Db      4921 TAGCCCTGATGTCTTGAAGAACTCCCTAACCCATGCTCCAGAAATTTGAGCAGTCTTC      4980
Qy      4981 CAATTTATATTTATCAAGAGACTGAGAAAGTACTCATGATAGCAGAGAGAGAGCGGCA      5040
Db      4981 CAATTTATATTTATCAAGAGACTGAGAAAGTACTCATGATAGCAGAGAGAGAGCGGCA      5040
Qy      5041 CTGTGCTTGTGAGCGTGAAGAAAGTGAACAGTCCCTGCGCCAGTCCCACTGCTGACC      5100
Db      5041 CTGTGCTTGTGAGCGTGAAGAAAGTGAACAGTCCCTGCGCCAGTCCCACTGCTGACC      5100
Qy      5101 CAGCCGACATCTCACCACCAATTTTGAAGCTGTCAAGGAGCTGCCATTTGTTGGGCA      5160
Db      5101 CAGCCGACATCTCACCACCAATTTTGAAGCTGTCAAGGAGCTGCCATTTGTTGGGCA      5160
Qy      5161 GCGAAGATTGAGAACGGGCTTGCATCTGTGACAGCCATGCGCAGCAAAAGTGTCTTC      5220
Db      5161 GCGAAGATTGAGAACGGGCTTGCATCTGTGACAGCCATGCGCAGCAAAAGTGTCTTC      5220
Qy      5221 CGCTACACGAAAACCTCAGAAATCTGCAATCCGGAAGAGATGAGAACTCAGAGCCC      5280
Db      5221 CGCTACACGAAAACCTCAGAAATCTGCAATCCGGAAGAGATGAGAACTCAGAGCCC      5280
Qy      5281 TGCAGCTGTATCACTTCAACCAATTAAGATGATCTGATGAAACCAATTAATTTAGCA      5340
Db      5281 TGCAGCTGTATCACTTCAACCAATTAAGATGATCTGATGAAACCAATTAATTTAGCA      5340
Qy      5341 ATGCATGAGAGAGTACAGCTCGAGGAATTCCTGGATGAAGATGACATTTCTGGCA      5400
Db      5341 ATGCATGAGAGAGTACAGCTCGAGGAATTCCTGGATGAAGATGACATTTCTGGCA      5400
Qy      5401 CTTGCTGTGTTGCGGCTCTTCAACAGCTTCCCTGTCTCAATCTGTCAGAGTGAACGC      5460
Db      5401 CTTGCTGTGTTGCGGCTCTTCAACAGCTTCCCTGTCTCAATCTGTCAGAGTGAACGC      5460
Qy      5461 GCAGGGCAGCAGAGAGATCTTGTGTGTTTCCAGAAATTTGAGAGTGTGTGTGATTTCT      5520
Db      5461 GCAGGGCAGCAGAGAGATCTTGTGTGTTTCCAGAAATTTGAGAGTGTGTGTGATTTCT      5520
Qy      5521 TACGGAAGACGTAGCCGACAGAGATCTCAAGTGAAGTTCCTTGGCTTTGGCC      5580

```

```

Db      5521 TACGGAAGACGTAGCCGACAGAGATCTCAAGTGAAGTTCCTTGGCTTTGGCC      5580
Qy      5581 TACAGAGAACCTATCTGTTTGTGACCCACTTCAACTCTGAGATTAATTGATTCAG      5640
Db      5581 TACAGAGAACCTATCTGTTTGTGACCCACTTCAACTCTGAGATTAATTGATTCAG      5640
Qy      5641 GCAGCTCTTCAAGAGGAACTTCCCTGCGAGCTGATCTGACATCTCCGAACTCCGCTAC      5700
Db      5641 GCAGCTCTTCAAGAGGAACTTCCCTGCGAGCTGATCTGACATCTCCGAACTCCGCTAC      5700
Qy      5701 CTGGGCGCTGCGCAATTTCTCAAGAGAGATTTACTTGGCGTCTCATACAGATAAATTA      5760
Db      5701 CTGGGCGCTGCGCAATTTCTCAAGAGAGATTTACTTGGCGTCTCATACAGATAAATTA      5760
Qy      5761 AGGATCATTTGTCTGCAAGAGAACTCTGTAAGAGATCCGCACTGAACCAACCGGCGC      5820
Db      5761 AGGATCATTTGTCTGCAAGAGAACTCTGTAAGAGATCCGCACTGAACCAACCGGCGC      5820
Qy      5821 CCGTCCACTTCCGACAGAGCCCAACAGAGAGCCCAACAGTCAACAGACATTC      5880
Db      5821 CCGTCCACTTCCGACAGAGCCCAACAGAGAGCCCAACAGTCAACAGACATTC      5880
Qy      5881 ACCAAGGCGTGGCTTCCAGCCCAAGGCGCGCCGGAAGGCGCCAGCCGAGAGCA      5940
Db      5881 ACCAAGGCGTGGCTTCCAGCCCAAGGCGCGCCGGAAGGCGCCAGCCGAGAGCA      5940
Qy      5941 AGCACAACCCACCGCTACCGCGAGGCGGAGCCGAGCTGCGAGAGGACAAATCTCTGAC      6000
Db      5941 AGCACAACCCACCGCTACCGCGAGGCGGAGCCGAGCTGCGAGAGGACAAATCTCTGAC      6000
Qy      6001 CGCCCTCTGAGCGAGAGAGTCCCGCGCGATCTCAGACGCGGAGAGAGCGGTCC      6060
Db      6001 CGCCCTCTGAGCGAGAGAGTCCCGCGCGATCTCAGACGCGGAGAGAGCGGTCC      6060
Qy      6061 CCGCGAGGCTGTTTGAAGACAGCAGCAGGCGCGGCTGCGGAGCGGTGAGACC      6120
Db      6061 CCGCGAGGCTGTTTGAAGACAGCAGCAGGCGCGGCTGCGGAGCGGTGAGACC      6120
Qy      6121 CCGCTGTCCAGGTGAACAGGTCTGAGAACCAATCTTCAATATA      6165
Db      6121 CCGCTGTCCAGGTGAACAGGTCTGAGAACCAATCTTCAATATA      6165

RESULT 3
US-10-415-011-43
: Sequence 43, Application US/10415011
: Publication No. US20040035394A1
: GENERAL INFORMATION:
: APPLICANT: INCYTE CORPORATION
: APPLICANT: GURURAJAN, Rajagopal
: APPLICANT: BAUGHN, Mariah R.
: APPLICANT: CHAWLA, Narinder K.
: APPLICANT: ELIOTT, Vicki S.
: APPLICANT: XU, Yuming
: APPLICANT: ARVIZU, Chandra S.
: APPLICANT: YAO, Monique G.
: APPLICANT: RAMKUMAR, Jayalaxmi
: APPLICANT: DING, Li
: APPLICANT: TANG, Y. Tom
: APPLICANT: HAFALIA, April J.A.
: APPLICANT: NGUYEN, Daniel B.
: APPLICANT: GANDHI, Ameena R.
: APPLICANT: LU, Van
: APPLICANT: YUE, Henry
: APPLICANT: BURFORD, Neil
: APPLICANT: BANDMAN, Olga
: APPLICANT: TRIBOULEY, Catherine M.
: APPLICANT: LAL, Preeti G.
: APPLICANT: RECIPION, Shirley A.
: APPLICANT: LU, Dying Aina M.
: APPLICANT: BOKOMSKY, Mark L.
: APPLICANT: THORNTON, Michael B.
: APPLICANT: SWARNAKER Anita

```


APPLICANT: THANGAVELU, Kavitha
APPLICANT: KHAN, Farrah A.
APPLICANT: ISON, Craig H.
TITLE OR INVENTION: HUMAN KINASES
FILE REFERENCE: PI-0262 USN
CURRENT APPLICATION NUMBER: US/10/415, 011
CURRENT FILING DATE: 2003-04-18
PRIORITY APPLICATION NUMBER: PCT/US01/47728
PRIORITY FILING DATE: 2001-10-20
PRIORITY APPLICATION NUMBER: US 60/242,410
PRIORITY FILING DATE: 2000-10-20
PRIORITY APPLICATION NUMBER: US 60/244,068
PRIORITY FILING DATE: 2000-10-27
PRIORITY APPLICATION NUMBER: US 60/245,708
PRIORITY FILING DATE: 2000-11-03
PRIORITY APPLICATION NUMBER: US 60/247,672
PRIORITY FILING DATE: 2000-11-09
PRIORITY APPLICATION NUMBER: US 60/249,565
PRIORITY FILING DATE: 2000-11-16
PRIORITY APPLICATION NUMBER: US 60/252,730
PRIORITY FILING DATE: 2000-11-22
PRIORITY APPLICATION NUMBER: US 60/250,807
PRIORITY FILING DATE: 2000-12-01
SOFTWARE: PERL Program
SEQ ID NO 43
LENGTH: 6298
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
OTHER INFORMATION: Incyte ID No. US20040053394A1 7484498CBI
US-10-415-011-43

Query Match 99.9%; Score 6161.8; DB 17; Length 6298;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 6163; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

1 ATGTGGAAGTTCAATATGAGAGCGGGAATCTTTGATGCTGCTGTAACCCATT 60
55 ATGTGGAAGTTCAATATGAGAGCGGGAATCTTTGATGCTGCTGTAACCCATT 114
61 GCCAGCGGCGCTCCAGGCTGAATCTGTCTCCAGGGAACACCCCTTATGACCTCA 120
115 GCCAGCGGCGCTCCAGGCTGAATCTGTCTCCAGGGAACACCCCTTATGACCTCA 174
121 CAGCAGATGTCCTCTCTTCCGAGAGGATATAGATCCCTCTTGTCTCTTTGAA 180
175 CAGCAGATGTCCTCTCTTCCGAGAGGATATAGATCCCTCTTGTCTCTTTGAA 234
181 GAATGCACTGACCTGCTCTGATGAAGATTAGCACTGAGCAACTTTGTCCGAAATAT 240
235 GAATGCACTGACCTGCTCTGATGAAGATTAGCACTGAGCAACTTTGTCCGAAATAT 294
241 TCCGACACCATAGCTAGTTAGAGAGGCTCCAGCTTCCGCAAGAGACTTCCAGATCAGA 300
295 TCCGACACCATAGCTAGTTAGAGAGGCTCCAGCTTCCGCAAGAGACTTCCAGATCAGA 354
301 AGTCTTGAGTGTGTGCTCACTTTGCTGAAGTCCAGGTGTTAGAGAGAAAGCAACCGGG 360
355 AGTCTTGAGTGTGTGCTCACTTTGCTGAAGTCCAGGTGTTAGAGAGAAAGCAACCGGG 414
361 GACATCTATGCTATGAAAGTGAAGAAAGAGGCTTTATTTGGCCAGAGGAGGTTTCA 420
415 GACATCTATGCTATGAAAGTGAAGAAAGAGGCTTTATTTGGCCAGAGGAGGTTTCA 474
421 TTTTGTGAGAGAGGAGGAAATATATCTCCAGAGCAAGCCCGTGGATCCCCCAATTA 480
475 TTTTGTGAGAGAGGAGGAAATATATCTCCAGAGCAAGCCCGTGGATCCCCCAATTA 534
481 CAGTATGCTTTTCAGAGCAAAATCACCTTTATCTGCTCATGGAATATCAGCTGGAGGG 540
535 CAGTATGCTTTTCAGAGCAAAATCACCTTTATCTGCTCATGGAATATCAGCTGGAGGG 594

541 GACTTGCTGTCACTTTGAAATAGATATAGAGCAAGTTAGTGAACCTTGATACAGTTT 600
595 GACTTGCTGTCACTTTGAAATAGATATAGAGCAAGTTAGTGAACCTTGATACAGTTT 654
601 TACCTAGCTGAGCGATTTTGGCTGTTCAAGCGCTTATCTGAGGGAATGCTGATCCGA 660
655 TACCTAGCTGAGCGATTTTGGCTGTTCAAGCGCTTATCTGAGGGAATGCTGATCCGA 714
661 GACATCAAGCCTGAGAACTTCTGTTGACCGACAGACATCAAGCTGTGGAATTTT 720
715 GACATCAAGCCTGAGAACTTCTGTTGACCGACAGACATCAAGCTGTGGAATTTT 774
721 GGAATCTCCCGGAAATGAAATTCAAACAAGATGTAATCCAACTCCGATTTGGAGC 780
775 GGAATCTCCCGGAAATGAAATTCAAACAAGATGTAATCCAACTCCGATTTGGAGC 834
781 CCGATTAACATGCTCTGTAAGTGTGTAAGTGAAGGGAATGGAAGGACCTTAC 840
835 CCGATTAACATGCTCTGTAAGTGTGTAAGTGAAGGGAATGGAAGGACCTTAC 894
841 GGCCTGGAAGTGAAGTGTGTAAGTGAAGGGAATGGAAGGACCTTAC 900
895 GGCCTGGAAGTGAAGTGTGTAAGTGAAGGGAATGGAAGGACCTTAC 954
901 TCCCTCTTCCAGAGGGAACCTTCCAGAACCTTCAATTAATTAATTAATTAATTAATTA 960
955 TCCCTCTTCCAGAGGGAACCTTCCAGAACCTTCAATTAATTAATTAATTAATTAATTA 1014
961 TTTTGAATTTTCCAGATGACCCCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1020
1015 TTTTGAATTTTCCAGATGACCCCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1074
1021 TTGTTGTCGCGCAGAAAG 1080
1075 TTGTTGTCGCGCAGAAAG 1134
1081 TCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1140
1135 TCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1194
1141 TCTGAAGATGACCTTCCAAATTTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1200
1195 TCTGAAGATGACCTTCCAAATTTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1254
1201 CCGTGCACAGTGAAGCTTCCAGGCTTCCGAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAG 1260
1255 CCGTGCACAGTGAAGCTTCCAGGCTTCCGAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAG 1314
1261 TACAGCAAGGCACTGGGGAATCTTGGTAGATCTGATCTGTGTGTGTCGGGCTTGAACCTCC 1320
1315 TACAGCAAGGCACTGGGGAATCTTGGTAGATCTGATCTGTGTGTGTCGGGCTTGAACCTCC 1374
1321 CCGTGCACAGTGAAGCTTCCAGGCTTCCGAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAG 1380
1375 CCGTGCACAGTGAAGCTTCCAGGCTTCCGAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAG 1434
1381 TCTCAAGCAAGTGTCAACAAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1440
1435 TCTCAAGCAAGTGTCAACAAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1494
1441 GAGGTGAAGGCTGTGCTTAGTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1500
1495 GAGGTGAAGGCTGTGCTTAGTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1554
1501 TCCCTCTTCCAG 1560
1555 TCCCTCTTCCAG 1614
1561 TTGAGAGCAAGCAGATGGAAGGTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1620
1615 TTGAGAGCAAGCAGATGGAAGGTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1674

QY 1621 GATATCAGAGACGAGCCGGAAGCTCCAGAAATCAAGAGCAGAGCTACAGGCTCAA 1680
Db 1675 GATATCAGAGACGAGCCGGAAGCTCCAGAAATCAAGAGCAGAGCTACAGGCTCAA 1734
QY 1681 GTGGAAGAAATGAGGTTGATGATGAATCAATTGGAAGAGATCTTGTCTCAGCAAGAA 1740
Db 1735 GTGGAAGAAATGAGGTTGATGATGAATCAATTGGAAGAGATCTTGTCTCAGCAAGAA 1794
QY 1741 CGAGTGAATCTTAAGAACTGAGCTGAGAGAGTCTCGGCTTGTCTGAAGAAATTCAG 1800
Db 1795 CGAGTGAATCTTAAGAACTGAGCTGAGAGAGTCTCGGCTTGTCTGAAGAAATTCAG 1854
QY 1801 CGGAAACGACAGAAATGTCAGCAATTAAGCTTTGAAGCTTAAGATCAAGGAAAGCTGAA 1860
Db 1855 CGGAAACGACAGAAATGTCAGCAATTAAGCTTTGAAGCTTAAGATCAAGGAAAGCTGAA 1914
QY 1861 GTGGAGAAATATGGAAGAACTGGAAGAGATCAATGCTGAGCAGAGCTCAAAATTCAGAG 1920
Db 1915 GTGGAGAAATATGGAAGAACTGGAAGAGATCAATGCTGAGCAGAGCTCAAAATTCAGAG 1974
QY 1921 CTCGAAGAAATCTGGAAGAGCTGTAAAGCCAGCAACGAGGCCACCGAGCTGCGAG 1980
Db 1975 CTCGAAGAAATCTGGAAGAGCTGTAAAGCCAGCAACGAGGCCACCGAGCTGCGAG 2034
QY 1981 AATATCCGCCAGGCAAGAGAGCCGAGCGAGGGAGCTGGAAGAGCTGCAGAACCGAGAG 2040
Db 2035 AATATCCGCCAGGCAAGAGAGCCGAGCGAGGGAGCTGGAAGAGCTGCAGAACCGAGAG 2094
QY 2041 GATCTTCTGAAGCAATCAGAAAGAAAGCTGGTGAAGCTGAGAAACCGCCGATCTCTG 2100
Db 2095 GATCTTCTGAAGCAATCAGAAAGAAAGCTGGTGAAGCTGAGAAACCGCCGATCTCTG 2154
QY 2101 GAGAAACAAGATTAAGAGACTAGAGACCATGAGCGTAGAGAAAAAGACTGAAGATGAC 2160
Db 2155 GAGAAACAAGATTAAGAGACTAGAGACCATGAGCGTAGAGAAAAAGACTGAAGATGAC 2214
QY 2161 ATTCAGACAAATATCCCAAGATCCAGATCCAGATGGCTGATAAATTTCTGAGCTGAAGAG 2220
Db 2215 ATTCAGACAAATATCCCAAGATCCAGATGGCTGATAAATTTCTGAGCTGAAGAG 2274
QY 2221 AAAATGCGGAGGCGCAAGTCTCAGCCGAGCACTGAGAGTGCCTGAAACAGAAAGAG 2280
Db 2275 AAAATGCGGAGGCGCAAGTCTCAGCCGAGCACTGAGAGTGCCTGAAACAGAAAGAG 2334
QY 2281 CAGCACTATGAGAAAGATTAAGTGTGGAACAATCAGATTAAGAAAGACTGAGCTGAC 2340
Db 2335 CAGCACTATGAGAAAGATTAAGTGTGGAACAATCAGATTAAGAAAGACTGAGCTGAC 2394
QY 2341 AAGGAGACACTGAGAAACATGATCAGAGACACGAGAGAGGCGCATGAGAGGGCAAA 2400
Db 2395 AAGGAGACACTGAGAAACATGATCAGAGACACGAGAGAGGCGCATGAGAGGGCAAA 2454
QY 2401 ATTTTCAGCCGAAACGAAGGCGATGATCAATGCTATGATTTCCAAAGATCAATCCCTGGA 2460
Db 2455 ATTTTCAGCCGAAACGAAGGCGATGATCAATGCTATGATTTCCAAAGATCAATCCCTGGA 2514
QY 2461 CAGAGATTGTGGAATGCTGTGAAGCCAAATAACTGTCAGCAAAATAGCAGCTTTTAAAC 2520
Db 2515 CAGAGATTGTGGAATGCTGTGAAGCCAAATAACTGTCAGCAAAATAGCAGCTTTTAAAC 2574
QY 2521 CAAAGGAACATGAAGGCGCCAAAGAAAGATGATTTCTGAATCTCAGGCAACAGAAATTTTAC 2580
Db 2575 CAAAGGAACATGAAGGCGCCAAAGAAAGATGATTTCTGAATCTCAGGCAACAGAAATTTTAC 2634
QY 2581 CTGAGAGACACAGGCTGGGAATTTGAGAGCCGAGAACGAAACTGAGAGACAGCTGAG 2640
Db 2635 CTGAGAGACACAGGCTGGGAATTTGAGAGCCGAGAACGAAACTGAGAGACAGCTGAG 2694
QY 2641 AAGATCCGCCACCAAGACACAGATGACAAAGATGGGCTGTGGAACCTGAGAGCAAGATTTG 2700
Db 2695 AAGATCCGCCACCAAGACACAGATGACAAAGATGGGCTGTGGAACCTGAGAGCAAGATTTG 2754
QY 2701 CGGAGGTCAGTCTTAGAGCAAGAGAGCAAGAACTGAGCTCAAGCGCCAGCTCAGAG 2760

Db 2755 CGGAGGTCAGTCTTAGAGCAAGAGAGCAAGAACTGAGAGCTCAAGCGCCAGCTCAGAG 2814
QY 2761 CTAAGCTCTCCCTGCAAGAGCCGAGTCAAGTTGAACAGCCCTGCAAGCTGCAACGAGCG 2820
Db 2815 CTAAGCTCTCCCTGCAAGAGCCGAGTCAAGTTGAACAGCCCTGCAAGCTGCAACGAGCG 2874
QY 2821 GCCCTGAGAGCCAGCTTCCGCAAGCAAGACAGAGCTGGAAGAGACCAAGCAAGAGCT 2880
Db 2875 GCCCTGAGAGCCAGCTTCCGCAAGCAAGACAGAGCTGGAAGAGACCAAGCAAGAGCT 2934
QY 2881 GAAGAGAGATCCAGGCACTCAGGCAATTAGAGATGAATTCAGCGCAAAATTTGATGCT 2940
Db 2935 GAAGAGAGATCCAGGCACTCAGGCAATTAGAGATGAATTCAGCGCAAAATTTGATGCT 2994
QY 2941 CTTGCTAACAGCTGTACTGTATATCAGAGACCTGAGAGAGCAGCTAAACAGCTGACGAG 3000
Db 2995 CTTGCTAACAGCTGTACTGTATATCAGAGACCTGAGAGAGCAGCTAAACAGCTGACGAG 3054
QY 3001 GACAAAGCTGAATCTCAACAAACAAACTTCTGTCAAACAACTGATGAGGCTTCT 3060
Db 3055 GACAAAGCTGAATCTCAACAAACAAACTTCTGTCAAACAACTGATGAGGCTTCT 3114
QY 3061 GCGCGCAACGACGAGATTTGTAACAATGCGAAGTGAAGTGAACAATCCGCGGGAGATC 3120
Db 3115 GCGCGCAACGACGAGATTTGTAACAATGCGAAGTGAAGTGAACAATCCGCGGGAGATC 3174
QY 3121 ACGGAACGAGATGACAGCTTACAGCCAGAGCAAGACGATGAGGCTGTGAACACAG 3180
Db 3175 ACGGAACGAGATGACAGCTTACAGCCAGAGCAAGACGATGAGGCTGTGAACACAG 3234
QY 3181 TGCACATGCTGAGAGAAACAGGTCATGATTTGAGAGCCCTTAACGATGAGCTCTAGAA 3240
Db 3235 TGCACATGCTGAGAGAAACAGGTCATGATTTGAGAGCCCTTAACGATGAGCTCTAGAA 3294
QY 3241 AAAGAGGCGAGTGGAGGCGCTGGAAGAGCTCCTGGGTATGGAATCCCAATTTGAG 3300
Db 3295 AAAGAGGCGAGTGGAGGCGCTGGAAGAGCTCCTGGGTATGGAATCCCAATTTGAG 3354
QY 3301 TGTGCGGTTGAGAGCTGAGAGGATCTGGAACACGAGAAACAGAGCAGGCGAGAGCC 3360
Db 3355 TGTGCGGTTGAGAGCTGAGAGGATCTGGAACACGAGAAACAGAGCAGGCGAGAGCC 3414
QY 3361 GATCAGCGGATCACCGAGTCTCGCCAGGTGTGAGGCTGAGATGAAGAGCACAGGCT 3420
Db 3415 GATCAGCGGATCACCGAGTCTCGCCAGGTGTGAGGCTGAGATGAAGAGCACAGGCT 3474
QY 3421 GAGATTTCTCGCTGCAAGAGGCTCTCAAGAGACGAAGCTGAAGGCCGAGAGGCTCTCT 3480
Db 3475 GAGATTTCTCGCTGCAAGAGGCTCTCAAGAGACGAAGCTGAAGGCCGAGAGGCTCTCT 3534
QY 3481 GACAACTCAATGACCTGGAAGAAACATGCTATGCTGGAATGAATGCCGGAAGCTTA 3540
Db 3535 GACAACTCAATGACCTGGAAGAAACATGCTATGCTGGAATGAATGCCGGAAGCTTA 3594
QY 3541 CAGCAAGAGCTGAGACTGACAGAGCTCAACAGAGGCTTCTGGAAGAGCAAGCCAA 3600
Db 3595 CAGCAAGAGCTGAGACTGACAGAGCTCAACAGAGGCTTCTGGAAGAGCAAGCCAA 3654
QY 3601 TTACAGCAGCAGATGACCTGACAGAAATCAATTTTCCGCTGACTCAAGGACTGCAA 3660
Db 3655 TTACAGCAGCAGATGACCTGACAGAAATCAATTTTCCGCTGACTCAAGGACTGCAA 3714
QY 3661 GAAGCTCTAGATCGGGCTGATCTTACGAAAGCAAGAAAGTGAAGCTTGAAGTATCAGCTG 3720
Db 3715 GAAGCTCTAGATCGGGCTGATCTTACGAAAGCAAGAAAGTGAAGCTTGAAGTATCAGCTG 3774
QY 3721 GAAGATTCAGGTTCTCTATATTCATGAAAGGTGAAATGGAAGCAGATTTCTCAA 3780
Db 3775 GAAGATTCAGGTTCTCTATATTCATGAAAGGTGAAATGGAAGCAGATTTCTCAA 3834
QY 3781 CAAACCAACTCATTTATTTTCTGAAGCCAAATGGAACCACTGCTAAAAAGAAAAG 3840

Dh 3835 CAACCAAACTCATTTGATTTTCTGCAAGCAAAAATGACCAACCTGCTAAAAAGAAAAAG 3894
Qy 3841 GTTCTCTGCGATACATAGAGCTGAGCTGGCCCTGAGAAAGAGAAAGCTCGCTGTGCA 3900
Db 3895 GTTCTCTGCGATACATAGAGCTGAGCTGGCCCTGAGAAAGAGAAAGCTCGCTGTGCA 3954
Qy 3901 GAGCTAGAGAAAGCCCTTCAGAGACCCGCGATGAGCTCCGGTCCGCCCGGAGAGAAAGCT 3960
Db 3955 GAGCTAGAGAAAGCCCTTCAGAGAGCCCGCATGAGCTCCGGTCCGCCCGGAGAGAAAGCT 4014
Qy 3961 GCCCAGCGCAAAAGCAAGGACCAACCCATCCAGCGCAGCAAGCCGAGGACAG 4020
Db 4015 GCCCAGCGCAAAAGCAAGGACCAACCCATCCAGCGCAGCAAGCCGAGGACAG 4074
Qy 4021 ATGCCCATATCTCCGATCTGTCGGTCCGACAGACCAAGCCAGTCCCATGAGCTGTG 4080
Db 4075 ATGCCCATATCTCCGATCTGTCGGTCCGACAGACCAAGCCAGTCCCATGAGCTGTG 4134
Qy 4081 GCCCGGCATCTCAGCCGAGAAAGAGTCTCACTCAGAGAAATTTAGTCGGCGCTT 4140
Db 4135 GCCCGGCATCTCAGCCGAGAAAGAGTCTCACTCAGAGAAATTTAGTCGGCGCTT 4194
Qy 4141 AAGGAACGATGCAACCAATATCTCAGCGATTCACGTAAGGACTGAACATCGAGCC 4200
Db 4195 AAGGAACGATGCAACCAATATCTCAGCGATTCACGTAAGGACTGAACATCGAGCC 4254
Qy 4201 ACAAAGTGTGTGTGTCTGAGATACGCTGCACTTTGAGCGCAGCAATCCAAATGCTC 4260
Db 4255 ACAAAGTGTGTGTGTCTGAGATACGCTGCACTTTGAGCGCAGCAATCCAAATGCTC 4314
Qy 4261 GAATGTAGAGTATGTGTCAACCCAAAGTGTCCAGCGCTTCCAGGCACTGTGGGCTTG 4320
Db 4315 GAATGTAGAGTATGTGTCAACCCAAAGTGTCCAGCGCTTCCAGGCACTGTGGGCTTG 4374
Qy 4321 CCTGCTGAATATGTCACACATTCACCGAGGCTTCTGCGGTGCAAAATGAACTCCCA 4380
Db 4375 CCTGCTGAATATGTCACACATTCACCGAGGCTTCTGCGGTGCAAAATGAACTCCCA 4434
Qy 4381 GGTCTCAGACCAAGAGCCGACAGAGACTTGTGCACTTGTGAGGGGTGATGAAAGTCCC 4440
Db 4435 GGTCTCAGACCAAGAGCCGACAGAGACTTGTGCACTTGTGAGGGGTGATGAAAGTCCC 4494
Qy 4441 AGGAATPAAGAAAGAGAGCAAGAGGTGGAGACAGAAAGTACATGTCTCTGGAGGATCA 4500
Db 4495 AGGAATPAAGAAAGAGAGCAAGAGGTGGAGACAGAAAGTACATGTCTCTGGAGGATCA 4554
Qy 4501 AAAGTCTCATTTATGACATGAGAGCAGAGAGCTGACAGAGCCCGGTGAGAAATTT 4560
Db 4555 AAAGTCTCATTTATGACATGAGAGCAGAGAGCTGACAGAGCCCGGTGAGAAATTT 4614
Qy 4561 GAGCTGTGCTTCCGACGAGGATGTATCTATTCAATGTGCGTGTGTCTCGAATTC 4620
Db 4615 GAGCTGTGCTTCCGACGAGGATGTATCTATTCAATGTGCGTGTGTCTCGAATTC 4674
Qy 4621 GCAAAATACAGCCAAAGCAGATGTCCCATACATCTGAAGATGAAATCTACCCGACACC 4680
Db 4675 GCAAAATACAGCCAAAGCAGATGTCCCATACATCTGAAGATGAAATCTACCCGACACC 4734
Qy 4681 ACCTGTGCGCCGAGAGAAACCTCTACTTGTAGCTCCGAGCTTCCCTGCAAAAGCGG 4740
Db 4735 ACCTGTGCGCCGAGAGAAACCTCTACTTGTAGCTCCGAGCTTCCCTGCAAAAGCGG 4794
Qy 4741 TGGGTACCGGCTTTAGAAATCAGTTGTTCGAGGTGGAGAGTTTCTAGGAAAAAGCAGAA 4800
Db 4795 TGGGTACCGGCTTTAGAAATCAGTTGTTCGAGGTGGAGAGTTTCTAGGAAAAAGCAGAA 4854
Qy 4801 GCTGATGCTAAATCTGTTTGGAAATCTCCCTGTCTAAATCTGAAGGTATGACCTTAGAC 4860
Db 4855 GCTGATGCTAAATCTGTTTGGAAATCTCCCTGTCTAAATCTGAAGGTATGACCTTAGAC 4914
Qy 4861 ATGAATGACGCTGCTTCAAGTACCAAGTGTGTGTGGGACCAAGAGAAAGGCTC 4920
Db 4915 ATGAATGACGCTGCTTCAAGTACCAAGTGTGTGTGGGACCAAGAGAAAGGCTC 4974

Qy 4921 TAGCCCTGAATGTCTTGAATAAACTCCCTAACCCATGTCCCGAGAAATTTGAGAGCTTTC 4980
Db 4975 TAGCCCTGAATGTCTTGAATAAACTCCCTAACCCATGTCCCGAGAAATTTGAGAGCTTTC 5034
Qy 4981 CAAATTTATATTTATCAAGAGACCTGAGAAAGCTACTCATGATAGAGAGAAAGCGGACA 5040
Db 5035 CAAATTTATATTTATCAAGAGACCTGAGAAAGCTACTCATGATAGAGAGAAAGCGGACA 5094
Qy 5041 CTGTGTCTTGTGACGTAAGAAAGTAAACAGTCCCTGTGCCAGTCCCACTGTGCC 5100
Db 5095 CTGTGTCTTGTGACGTAAGAAAGTAAACAGTCCCTGTGCCAGTCCCACTGTGCC 5154
Qy 5101 CAGCCGACATCTCACCCCAATTTTGAAGCTGTCAAGGGCTGCCACTTGTGGGGCA 5160
Db 5155 CAGCCGACATCTCACCCCAATTTTGAAGCTGTCAAGGGCTGCCACTTGTGGGGCA 5214
Qy 5161 GCGAAGTTGAGAAAGCGGCTCTGCACTGTGTGACGCCATGCCCAAGATCTCATCTC 5220
Db 5215 GCGAAGTTGAGAAAGCGGCTCTGCACTGTGTGACGCCATGCCCAAGATCTCATCTC 5274
Qy 5221 CGCTAACAGAAACCTCAGCAATATCTGCATCCGAGAAAGATAGAGACTCAGAGCCC 5280
Db 5275 CGCTAACAGAAACCTCAGCAATATCTGCATCCGAGAAAGATAGAGACTCAGAGCCC 5334
Qy 5281 TGAGCTGTATTCACATTCACCAATTAACGTATCTCATTTGAAACCAATTAATCTAGAA 5340
Db 5335 TGAGCTGTATTCACATTCACCAATTAACGTATCTCATTTGAAACCAATTAATCTAGAA 5394
Qy 5341 ATGACATGAAAGAGATACAGCTCTGAGAAATCTCTGATTAAGATACCATCTCTTGCA 5400
Db 5395 ATGACATGAAAGAGATACAGCTCTGAGAAATCTCTGATTAAGATACCAATCTCTTGCA 5454
Qy 5401 CCTGCTGTGTTTCCGCCCTCTTCCAAAGCTTCCCTGTCTCAATTCGACAGTGAACAGC 5460
Db 5455 CCTGCTGTGTTTCCGCCCTCTTCCAAAGCTTCCCTGTCTCAATTCGACAGTGAACAGC 5514
Qy 5461 GCAGGGCAGCAGAGAGATACTTGTGTGTTCAGAAATTTGAGATGTTCTGTGATTC 5520
Db 5515 GCAGGGCAGCAGAGAGATACTTGTGTGTTCAGAAATTTGAGATGTTCTGTGATTC 5574
Qy 5521 TACGAAAGATGATAGCCGACACACATCTCAAGTGAATGCTTACTTGTGCTTTGCC 5580
Db 5575 TACGAAAGATGATAGCCGACACACATCTCAAGTGAATGCTTACTTGTGCTTTGCC 5634
Qy 5581 TACAGAGAACCCATATCTGTTTGTGACCACTTCAACTCTCGAAGTAAATGATCCAG 5640
Db 5635 TACAGAGAACCCATATCTGTTTGTGACCACTTCAACTCTCGAAGTAAATGATCCAG 5694
Qy 5641 GCAAGCTCTCAGCAGAGACCCCTGTGCCAGAGGTACCTGAGCAATCCGAAACCCGCGCTAC 5700
Db 5695 GCAAGCTCTCAGCAGAGACCCCTGTGCCAGAGGTACCTGAGCAATCCGAAACCCGCGCTAC 5754
Qy 5701 CTGGGCCCTGCAATTTCTCAGAGAGATTTACTTGTGCTCTCATACAGATTAATTA 5760
Db 5755 CTGGGCCCTGCAATTTCTCAGAGAGATTTACTTGTGCTCTCATACAGATTAATTA 5814
Qy 5761 AGGGTCAATTTGTGTCAGAGGAAACCTCGTGAAGAGTCCGAGCTGGAACCAACGAGGGC 5820
Db 5815 AGGGTCAATTTGTGTCAGAGGAAACCTCGTGAAGAGTCCGAGCTGGAACCAACGAGGGC 5874
Qy 5821 CCGTCAACTTCCGAGCAGAGCCCAACAGCAGGCCCCACCACTACACAGACGACATC 5880
Db 5875 CCGTCAACTTCCGAGCAGAGCCCAACAGCAGGCCCCACCACTACACAGACGACATC 5934
Qy 5881 ACCAAGCGGTGGCTTCCAGCCAGCGCGCCGAAAGGCCCAAGCCACCCGAGAGCCA 5940
Db 5935 ACCAAGCGGTGGCTTCCAGCCAGCGCGCCGAAAGGCCCAAGCCACCCGAGAGCCA 5994
Qy 5941 AGCACACCCACCGCTACCGAGGGGGGAGACCGAGCTGCGAGGGAACAAGTCTCTGGC 6000
Db 5995 AGCACACCCACCGCTACCGAGGGGGGAGACCGAGCTGCGAGGGAACAAGTCTCTGGC 6054

Qy 6001 CGCCCCCTGGAGCAAGAAAGTCCCCGGCCGGATCTCAGACAGCGGAGAGCGGTCC 6060
Db 6055 CGCCCCCTGGAGCAAGAAAGTCCCCGGCCGGATCTCAGACAGCGGAGAGCGGTCC 6114
Qy 6061 CCGCGAGGCTGTTTGAAGACAGACAGAGGGCCGGCTGCTGGGAGCCGTGAGGACC 6120
Db 6115 CCGCGAGGCTGTTTGAAGACAGACAGAGGGCCGGCTGCTGGGAGCCGTGAGGACC 6174
Qy 6121 CCGCTGCTCCAGGTGAACAAGTCTGGACCAAGTCTTCAATATA 6165
Db 6175 CCGCTGCTCCAGGTGAACAAGTCTGGACCAAGTCTTCAATATA 6219

RESULT 4
US-10-618-941-1
; Sequence 1, Application US/10618941
; Publication No. US20040197792A1
; GENERAL INFORMATION:
; APPLICANT: WHITE, DAVID
; APPLICANT: MANNING, GERARD
; APPLICANT: CANEPEL, SEAN
; TITLE OF INVENTION: NOVEL KINASES
; FILE REFERENCE: 034536-0321
; CURRENT APPLICATION NUMBER: US/10/618, 941
; CURRENT FILING DATE: 2003-07-15
; PRIOR APPLICATION NUMBER: 60/395, 632
; PRIOR FILING DATE: 2002-07-15
; NUMBER OF SEQ ID NOS: 143
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1
; LENGTH: 8656
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-618-941-1

Query Match 99.7%; Score 6144; DB 18; Length 8656;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 6160; Conservative 0; Mismatches 5; Indels 3; Gaps 1;

Qy 1 ATGTTGAAGTTCAAAATATGAGGCGGGAATCCTTGTGATGCTGGTGTGAACCCATT 60
Db 51 ATGTTGAAGTTCAAAATATGAGGCGGGAATCCTTGTGATGCTGGTGTGAACCCATT 110
Qy 61 GCCAGCGGCGCTCCAGGCTGAATCTGTTCTTCCAGGAGAAACCACTTATGACTCAA 120
Db 111 GCCAGCGGCGCTCCAGGCTGAATCTGTTCTTCCAGGAGAAACCACTTATGACTCAA 170
Qy 121 CAGCAGATGCTCTCTCTTCCGAGAAAGGATATTAAGTCCCTTTGTTCTCTTTGAA 180
Db 171 CAGCAGATGCTCTCTCTTCCGAGAAAGGATATTAAGTCCCTTTGTTCTCTTTGAA 230
Qy 181 GAATGAGTGAAGCTGCTGATGAAGATTAAGACGCTGACACTTTGCCGAA---G 237
Db 231 GAATGAGTGAAGCTGCTGATGAAGATTAAGACGCTGACACTTTGCCGAA---G 290
Qy 238 TATTCGACACCATAGCTAGATTACAGAGCTCCAGCTTGGGAAAGACTTGGAGTCC 297
Db 291 TATTCGACACCATAGCTAGATTACAGAGCTCCAGCTTGGGAAAGACTTGGAGTCC 350
Qy 298 AGAAGTCTTGTAGTGTGTGCTCACTTGTCTGAAGTGAAGTGTGAAGAGAAAGCAAC 357
Db 351 AGAAGTCTTGTAGTGTGTGCTCACTTGTCTGAAGTGAAGTGTGAAGAGAAAGCAAC 410
Qy 358 GGGGACATCTATGCTATGAAGATGAAGAAAGGCTTATTTGGCCCGAGAGCAAGT 417
Db 411 GGGGACATCTATGCTATGAAGATGAAGAAAGGCTTATTTGGCCCGAGAGCAAGT 470
Qy 418 TCAATTTTGAAGAGAGCGAACAATATATCTGAAAGCAACAAGCCGTGATCCCCCA 477
Db 471 TCAATTTTGAAGAGAGCGAACAATATATCTGAAAGCAACAAGCCGTGATCCCCCA 530
Qy 478 TTACAGTATGCTTTCAAGACAAAAATCACTTTATCTGATCATGAATATCAAGCTTGA 537
Db 537 TTACAGTATGCTTTCAAGACAAAAATCACTTTATCTGATCATGAATATCAAGCTTGA

Db 531 TTACAGTATGCTTTCAAGACAAAAATCACTTTATCTGATCATGAATATCAAGCTTGA 590
Qy 538 GGGGACATCTGCTCACTTTGAATAGATAGAGACCAATTAGATGAACCTGATACAG 597
Db 591 GGGGACATCTGCTCACTTTGAATAGATAGAGACCAATTAGATGAACCTGATACAG 650
Qy 598 TTTTACTTACTGAGCTGATTTTGGCTTTCAACAGCTTCACTGATGGGATAGTGCAT 657
Db 651 TTTTACTTACTGAGCTGATTTTGGCTTTCAACAGCTTCACTGATGGGATAGTGCAT 710
Qy 658 CGAGACATCAAGCCTGAGCAATTTCTGTTGACCGGACAGACATCAAGCTGTTGAT 717
Db 711 CGAGACATCAAGCCTGAGCAATTTCTGTTGACCGGACAGACATCAAGCTGTTGAT 770
Qy 718 TTTGATCTGCGCGAAATGAATTCACACAGATGATGCAACTCCGATTTGG 777
Db 771 TTTGATCTGCGCGAAATGAATTCACACAGATGATGCAACTCCGATTTGG 830
Qy 778 ACCCGAGTTACATGCTCTGAAAGTCTGATGATGAACGGGATGGAAGAGCAC 837
Db 831 ACCCGAGTTACATGCTCTGAAAGTCTGATGATGAACGGGATGGAAGAGCAC 890
Qy 838 TACGGCCTGAGCTGTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGAT 897
Db 891 TACGGCCTGAGCTGTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGAT 950
Qy 898 AGATCCCCCTTGGAGAGGAACTTGCAGAACCTTCAATACATTAATGAATTTTCAG 957
Db 951 AGATCCCCCTTGGAGAGGAACTTGCAGAACCTTCAATACATTAATGAATTTTCAG 1010
Qy 958 CGGTTTGAATTTTCCAGATGACCCCAAGTGAAGCTTCTTGTATCTGATTTCAA 1017
Db 1011 CGGTTTGAATTTTCCAGATGACCCCAAGTGAAGCTTCTTGTATCTGATTTCAA 1070
Qy 1018 AGCTGTGTGCGGCGCAAGAGAGACCTGAAGTTGAAGTCTTGTGATCTGATTTTC 1077
Db 1071 AGCTGTGTGCGGCGCAAGAGAGACCTGAAGTTGAAGTCTTGTGATCTGATTTTC 1130
Qy 1078 TTTCTTAAATTAAGTGAACCAATCTGTAATCTTCTCCCTTCTGTTCCACCTTC 1137
Db 1131 TTTCTTAAATTAAGTGAACCAATCTGTAATCTTCTCCCTTCTGTTCCACCTTC 1190
Qy 1138 AAGCTGACGATGACACCTTCAATTTTGAATGAACCAAGAAATGCTGGGTTTCATCC 1197
Db 1191 AAGCTGACGATGACACCTTCAATTTTGAATGAACCAAGAAATGCTGGGTTTCATCC 1250
Qy 1198 TCTCCGTCGACGCTGAGCCCTCAAGCTTCTCGGTTGAAGAACTGCTTGGGGTTT 1257
Db 1251 TCTCCGTCGACGCTGAGCCCTCAAGCTTCTCGGTTGAAGAACTGCTTGGGGTTT 1310
Qy 1258 TCGTACAGCAAGGCACTGGGAAATTTCTGTAGATCTGATCTGTTGTGCTGGATC 1317
Db 1311 TCGTACAGCAAGGCACTGGGAAATTTCTGTAGATCTGATCTGTTGTGCTGGATC 1370
Qy 1318 TCCCTGCAAGACTAGCTCCATGAAAGAACTTCTCATCAAAAGCAAGACTACAA 1377
Db 1371 TCCCTGCAAGACTAGCTCCATGAAAGAACTTCTCATCAAAAGCAAGACTACAA 1430
Qy 1378 GACTCTACAGACAGTGTCAAGATGAGACAGAAATGACCCGGTTACATCGAGAGTG 1437
Db 1431 GACTCTACAGACAGTGTCAAGATGAGACAGAAATGACCCGGTTACATCGAGAGTG 1490
Qy 1438 TCAAGGTGAGAGGCTGTAGTACAGAGAGGATGAGCTGAAAGGCTCTGAGACTCAG 1497
Db 1491 TCAAGGTGAGAGGCTGTAGTACAGAGAGGATGAGCTGAAAGGCTCTGAGACTCAG 1550
Qy 1498 AGATCCCTCTGAGACAGACTTGTCTACTACATCAAGATGCAATGACTTAAAGCGA 1557
Db 1551 AGATCCCTCTGAGACAGACTTGTCTACTACATCAAGATGCAATGACTTAAAGCGA 1610
Qy 1558 AGTTGAGCAAGCAAGATGAGAGTGTCCAGAGAGATGACAAAGCACTGACCTTCTC 1617
Db 1611 AGTTGAGCAAGCAAGATGAGAGTGTCCAGAGAGATGACAAAGCACTGACCTTCTC 1670

QY	1618	CATGATATCAGAGACAGAGCCGGAGCTTCAAGAAATCAAGACGAGTACCAAGCT	1677
Db	1671	CATGATATCAGAGACAGAGCCGGAGCTTCAAGAAATCAAGACGAGTACCAAGCT	1730
QY	1678	CAGTGGAAAGAAATGAGCTTGATGATGAATTCAGTTGGAAAGAGATCTTGTCTCAGCAAGA	1737
Db	1731	CAGTGGAAAGAAATGAGCTTGATGATGAATTCAGTTGGAAAGAGATCTTGTCTCAGCAAGA	1790
QY	1738	AGACGAGATGATCTTACGAATCTTGAGCTGAGAGAGTCTCGGCTTGTCTGGAAGATTC	1797
Db	1791	AGACGAGATGATCTTACGAATCTTGAGCTGAGAGAGTCTCGGCTTGTCTGGAAGATTC	1850
QY	1798	AAGCGGAAACCGACAGATGTGCACATTAACCTTTGAAGGCTTAAGATCAAGGAGGCT	1857
Db	1851	AAGCGGAAACCGACAGATGTGCACATTAACCTTTGAAGGCTTAAGATCAAGGAGGCT	1910
QY	1858	GAAATGGGGAATATGTCGAAACTGGAAGATCAATTCGTGAGCAGACACTCAAAATTCAG	1917
Db	1911	GAAATGGGGAATATGTCGAAACTGGAAGATCAATTCGTGAGCAGACACTCAAAATTCAG	1970
QY	1918	GAGCTTCAAGAGAACTGGAAGAGCTGTAAAGCCACAGACGAGGACCCGAGCTGCTG	1977
Db	1971	GAGCTTCAAGAGAACTGGAAGAGCTGTAAAGCCACAGACGAGGACCCGAGCTGCTG	2030
QY	1978	CAGATATTCCTCCGAGCAAGAGCGAGCCGAGAGGAGCTTGAGAACTCTCAGAACCGA	2037
Db	2031	CAGATATTCCTCCGAGCAAGAGCGAGCCGAGAGGAGCTTGAGAACTCTCAGAACCGA	2090
QY	2038	GAGATATTCCTGAAAGGCATCAGAAAGAAAGCTGTGGAAGCTGAGAAACGCGGCATTC	2097
Db	2091	GAGATATTCCTGAAAGGCATCAGAAAGAAAGCTGTGGAAGCTGAGAAACGCGGCATTC	2150
QY	2098	CTGGAGAACCAAGTAAAGAGCTAGAGAACAATGAGCCTAGAGAAACAGACTGAAGGAT	2157
Db	2151	CTGGAGAACCAAGTAAAGAGCTAGAGAACAATGAGCCTAGAGAAACAGACTGAAGGAT	2210
QY	2158	GACATCCGAGCAAAATCCCAACAGATCCAGCAGATGCTGATTAATTTCTGAGCTGAA	2217
Db	2211	GACATCCGAGCAAAATCCCAACAGATCCAGCAGATGCTGATTAATTTCTGAGCTGAA	2270
QY	2218	GAGAAACATCGGAGGAGCCCAAGTCTCAACCCAGCACCTAGAGAGTGCACCTGAACAGAAA	2277
Db	2271	GAGAAACATCGGAGGAGCCCAAGTCTCAACCCAGCACCTAGAGAGTGCACCTGAACAGAAA	2330
QY	2278	GAGCAGCACTATGAGAAAAGATTAAAGTTTGACAAATCAGATTAAGAAAAGACTGCT	2337
Db	2331	GAGCAGCACTATGAGAAAAGATTAAAGTTTGACAAATCAGATTAAGAAAAGACTGCT	2390
QY	2338	GACAAGGAGACACTGAGAAACATGATGAGAGACACGAGGAGGAGGCCCATGAGAAAGGC	2397
Db	2391	GACAAGGAGACACTGAGAAACATGATGAGAGACACGAGGAGGAGGCCCATGAGAAAGGC	2450
QY	2398	AAAAATTCACGCGAACAGAGGCGCATGATCATGATTCAGATTCAGATTCAGATTCCTG	2457
Db	2451	AAAAATTCACGCGAACAGAGGCGCATGATCATGATTCAGATTCAGATTCAGATTCCTG	2510
QY	2458	GAAACGAGAGTTGTGGAACCTGTCTGAGCCATTAACCTTGACGAAATAGCAGCTTTT	2517
Db	2511	GAAACGAGAGTTGTGGAACCTGTCTGAGCCATTAACCTTGACGAAATAGCAGCTTTT	2570
QY	2518	AACCCAAAGAACATGAGAGGCCCAAGAGAGATGATTTCTGAACCTCAGGCAACAGAAATTT	2577
Db	2571	AACCCAAAGAACATGAGAGGCCCAAGAGAGATGATTTCTGAACCTCAGGCAACAGAAATTT	2630
QY	2578	TACCTGAGACACAGGCTGGAGAGTTGAGAGCCACGAAACCGAAACTTGAGAGAGCTG	2637
Db	2631	TACCTGAGACACAGGCTGGAGAGTTGAGAGCCACGAAACCGAAACTTGAGAGAGCTG	2690
QY	2638	GAGAAAGATCAGACCAACAGACCAACATGAGCTGCTGGAATCTGAGACAGACA	2697
Db	2691	GAGAAAGATCAGACCAACAGACCAACATGAGCTGCTGGAATCTGAGACAGACA	2750

QY	2698	TTGCGGAGGCTCAGTCTTGAAGCA	CGAGAGCA	GAAATCTG	AGCTCA	AGCCCA	AGTCA	CA	2752
Db	2751	TTGCGGAGGCTCAGTCTTGAAGCA	CGAGAGCA	GAAATCTG	AGCTCA	AGCCCA	AGTCA	CA	2810
QY	2758	GAGCTACAGCTCTCCCTCGAGGAG	GCGAGTCA	CAAGTTGA	CAAGCCCTG	CAAGCTG	CAAGG		2817
Db	2811	GAGCTACAGCTCTCCCTCGAGGAG	GCGAGTCA	CAAGTTGA	CAAGCCCTG	CAAGCTG	CAAGG		2870
QY	2818	GCGGCGCTGAGAGGCGAGCTTTCG	CAAGGCGA	AGACAGAGCTG	GAAGAGCA	CAAGAGAA		2877	
Db	2871	GCGGCGCTGAGAGGCGAGCTTTCG	CAAGGCGA	AGACAGAGCTG	GAAGAGCA	CAAGAGAA		2930	
QY	2878	GCTGAAGAGAGATTCAGGCACTCA	CGGCAATG	AGATGAATTC	ACGCGAAATTTGAT			2937	
Db	2931	GCTGAAGAGAGATTCAGGCACTCA	CGGCAATG	AGATGAATTC	ACGCGAAATTTGAT			2990	
QY	2938	GCTCTTTCTTAACGCTGTACTGTAT	TCAACA	CTTGAGGAGGAGCTTAA	CCACTGACC			2997	
Db	2991	GCTCTTTCTTAACGCTGTACTGTAT	TCAACA	CTTGAGGAGGAGCTTAA	CCACTGACC			3050	
QY	2998	GAGACAAAGCTGAATCAACAA	CCAAACTTTACTG	TGCACAAACA	CTGATGAGGCT			3057	
Db	3051	GAGACAAAGCTGTAATCAACAA	CCAAACTTTACTG	TGCACAAACA	CTGATGAGGCT			3110	
QY	3058	TCTGGCGCCAAAGAGATTGTACAA	CTCGAAGTGA	ATGATGACCA	TCTCCGCGGAG			3117	
Db	3111	TCTGGCGCCAAAGAGATTGTACAA	CTCGAAGTGA	ATGATGACCA	TCTCCGCGGAG			3170	
QY	3118	ATCAGGAAACGAGAGATGCAAGCTT	ACCA	CGCAGAACAA	ACGATGAGGCT	CTGAAAGC		3177	
Db	3171	ATCAGGAAACGAGAGATGCAAGCTT	ACCA	CGCAGAACAA	ACGATGAGGCT	CTGAAAGC		3230	
QY	3178	ACGTGCACATGCTGAGAGGAA	CAGGTCA	TGGAAGGCGCCTTAA	CGATGAGCTGTA			3237	
Db	3231	ACGTGCACATGCTGAGAGGAA	CAGGTCA	TGGAAGGCGCCTTAA	CGATGAGCTGTA			3290	
QY	3238	GAAAAAGACGCGCATGTGGAGG	CGCTGGA	AGACGCTCTCGGCTGTATGA	AAATCCCA	ATT		3297	
Db	3291	GAAAAAGACGCGCATGTGGAGG	CGCTGGA	AGACGCTCTCGGCTGTATGA	AAATCCCA	ATT		3350	
QY	3298	GAGTGTGCGGTTTCAAGAGCTG	CAAGAGATG	CTGGAAC	CCGAGAA	CAAGACGAGGCGAGA		3357	
Db	3351	GAGTGTGCGGTTTCAAGAGCTG	CAAGAGATG	CTGGAAC	CCGAGAA	CAAGACGAGGCGAGA		3410	
QY	3358	GCCCATCAGCGGATCA	CCGAGTCTCG	CGAGGTGGA	GCTGCGACAG	AGAGAGGACCAAG		3417	
Db	3411	GCCCATCAGCGGATCA	CCGAGTCTCG	CGAGGTGGA	GCTGCGACAG	AGAGAGGACCAAG		3470	
QY	3418	GCTGAGATTCTCGCTCTG	CAGAGGCTCTCA	AAAGACGAGAGCTGA	AGCCGAGAGCTTC			3477	
Db	3471	GCTGAGATTCTCGCTCTG	CAGAGGCTCTCA	AAAGACGAGAGCTGA	AGCCGAGAGCTTC			3530	
QY	3478	TCTGACAGCTCAATGA	ACTTGA	GAGAGAGACATG	CTATGCTTGA	ATGAATGAAATG	ATGATG	3537	
Db	3531	TCTGACAGCTCAATGA	ACTTGA	GAGAGAGACATG	CTATGCTTGA	ATGAATGAAATG	ATGATG	3590	
QY	3538	TTTACAGAGAAAGCTG	GAGACTGAA	CCGAGACTCA	AAACAGAGGCTT	CTGAAAGGCAAGCC		3597	
Db	3591	TTTACAGAGAAAGCTG	GAGACTGAA	CCGAGACTCA	AAACAGAGGCTT	CTGAAAGGCAAGCC		3650	
QY	3598	AAATTTACAGACGAGATG	AGACTTGC	AGAAAAATCA	CAATTTCCGTCTGA	CTCAAGAGCTG		3657	
Db	3651	AAATTTACAGACGAGATG	AGACTTGC	AGAAAAATCA	CAATTTCCGTCTGA	CTCAAGAGCTG		3710	
QY	3658	CAAGAGCTCTAAGTCGGGCTG	ATCTTA	CTGAGACAGAA	GAAGATGAC	TTGAGATTCAG		3717	
Db	3711	CAAGAGCTCTAAGTCGGGCTG	ATCTTA	CTGAGACAGAA	GAAGATGAC	TTGAGATTCAG		3770	
QY	3718	CTGAAAAACATTCAGGTTCT	CTCTAT	TTCTCATGAA	AAAGTGA	ATGAAATGGA	AGCACTAT	3777	
Db	3771	CTGAAAAACATTCAGGTTCT	CTCTAT	TTCTCATGAA	AAAGTGA	ATGAAATGGA	AGCACTAT	3830	
QY	3778	CAAGCAACCAAACTCATATG	ATTTTCTT	CTGCAAGCCAA	AAATGGA	CCAACTCTGTAA	AAAAAGAA	3837	

D	3831	CAACAAACAACTCATTTGATTTTCTGACAGCCAAATGACCAACTGCTAAAAAGAA	3890	D	4911	GACATGAACCTGACAGCTGCCCTTCACTGACAGAGTGTGTTGGTGGGCACCGAGAAAGG	4970
Q	3838	AAGGTCCTCTGAGTACAAATGAGCTGAGCTGGCCCTGAGAGAGAAAGCTCGCTGT	3897	Q	4918	CTCTAGCCCTGAAATGTTTGAAGAACTCCCTAACCATATGCCAGAAATTTGAGCAATC	4977
D	3891	AAGGTCCTCTGAGTACAAATGAGCTGAGCTGGCCCTGAGAGAGAAAGCTCGCTGT	3950	D	4971	CTCTAGCCCTGAAATGTTTGAAGAACTCCCTAACCATATGCCAGAAATTTGAGCAATC	5030
Q	3898	GCAAGCTAGAGAGAGCCCTTCAAGAGACCCGATCGAGCTCCGCTCCGCCGAGAGAA	3957	Q	4978	TTCCAAATTTATTTATCAAGAGACTGAGAACTACTCATGATAGAGAGAGAAAGACCG	5037
D	3951	GCAAGCTAGAGAGAGCCCTTCAAGAGACCCGATCGAGCTCCGCTCCGCCGAGAGAA	4010	D	5031	TTCCAAATTTATTTATCAAGAGACTGAGAACTACTCATGATAGAGAGAGAAAGACCG	5090
Q	3958	GCTGGCCACCGCAAGAGACGACCAACCAACCAACCAACCAACCAACCAACCAACCA	4017	Q	5038	GCACTGTGCTTGGAGCGTGAAGAAAGTAAACAGTCCCTGGCCAGTCCCACTGCT	5097
D	4011	GCTGGCCACCGCAAGAGACGACCAACCAACCAACCAACCAACCAACCAACCAACCA	4070	D	5091	GCACTGTGCTTGGAGCGTGAAGAAAGTAAACAGTCCCTGGCCAGTCCCACTGCT	5150
Q	4018	CAGATCGCATGTCCTGCGATCGTGGTCCGACAGAGACAGCCAGTGCATGAGCTG	4077	Q	5098	GCCGAGCCGACATCTCACCCAACTTTTGAAGCTGTCAAGAGGCTGCCACTGTTGGG	5157
D	4071	CAGATCGCATGTCCTGCGATCGTGGTCCGACAGAGACAGCCAGTGCATGAGCTG	4130	D	5151	GCCGAGCCGACATCTCACCCAACTTTTGAAGCTGTCAAGAGGCTGCCACTGTTGGG	5210
Q	4078	CTGGCCCGGCATCTCAGCCGCAAGAGAGTCTTCACTCCAGAGAAATTTAGTCCGCT	4137	Q	5158	GCAAGGCAAGTTGAGAAAGGCGCTCTGATCTGTGCAACCATATGCCAGCAAGTGTCT	5217
D	4131	CTGGCCCGGCATCTCAGCCGCAAGAGAGTCTTCACTCCAGAGAAATTTAGTCCGCT	4190	D	5211	GCAAGGCAAGTTGAGAAAGGCGCTCTGATCTGTGCAACCATATGCCAGCAAGTGTCT	5270
Q	4138	CTTAAAGAAAGCATGACACCAATATTTCTCAACGATTTCAACGATGAGACTGACATG	4197	Q	5218	CTCCGCTACAGAGAAACCTCAGCAATACTGATCCGGAAGATAGAGACCTCAGAG	5277
D	4191	CTTAAAGAAAGCATGACACCAATATTTCTCAACGATTTCAACGATGAGACTGACATG	4250	D	5271	CTCCGCTACAGAGAAACCTCAGCAATACTGATCCGGAAGATAGAGACCTCAGAG	5330
Q	4198	GCCACAAAGT	4257	Q	5278	CCCTGACGCTGTATTCATCTTCAACCAATTAAGTATCTCTTGAAGCAATTAATTTAC	5337
D	4251	GCCACAAAGT	4310	D	5331	CCCTGACGCTGTATTCATCTTCAACCAATTAAGTATCTCTTGAAGCAATTAATTTAC	5390
Q	4258	CTGGAATGTGAGT	4317	Q	5338	GAATGCAATGAGAGAGTACAGCTCTGAGAAATTCCTGTGATTAAGATGACATTCCTG	5397
D	4311	CTGGAATGTGAGT	4370	D	5391	GAATGCAATGAGAGAGTACAGCTCTGAGAAATTCCTGTGATTAAGATGACATTCCTG	5450
Q	4318	TTGGCTGCTGAATATGCTCAACACTTACCGAGGCTTTGCGGTGAACAAATGAACTCC	4377	Q	5398	GCACTGTGTGTTTGGCGGCTCTTCCAAACGCTTCCCTCTCAATCGGAGGAGTAAAC	5457
D	4371	TTGGCTGCTGAATATGCTCAACACTTACCGAGGCTTTGCGGTGAACAAATGAACTCC	4430	D	5451	GCACTGTGTGTTTGGCGGCTCTTCCAAACGCTTCCCTCTCAATCGGAGGAGTAAAC	5510
Q	4378	CGAGGTCTCAGACCAAGAGAGCCAGAGAGCTTGCACCTGAGAGGAGTGAATGAAGTG	4437	Q	5458	AGCGAGGGCAGAGAGAGAGTACTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	5517
D	4431	CGAGGTCTCAGACCAAGAGAGCCAGAGAGCTTGCACCTGAGAGGAGTGAATGAAGTG	4490	D	5511	AGCGAGGGCAGAGAGAGAGTACTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	5570
Q	4438	CCGAGAAATTAACAAACGAGACAGCAAGGCTGGAGAGAGAGTACATTTGCTTGGAGGA	4497	Q	5518	TCTTACGAAAGAGTACCGGCAACAGATCTCAAGTGAAGTGTGCTTACCTTTGGCTTT	5577
D	4491	CCGAGAAATTAACAAACGAGACAGCAAGGCTGGAGAGAGAGTACATTTGCTTGGAGGA	4550	D	5571	TCTTACGAAAGAGTACCGGCAACAGATCTCAAGTGAAGTGTGCTTACCTTTGGCTTT	5630
Q	4498	TCAAAAGTCTCTATTATGACATGAAGCAAGAGACTGAGACAGAGCCGCTGAGAGAA	4557	Q	5578	GCTTACGAGAAACCTTATGTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	5637
D	4551	TCAAAAGTCTCTATTATGACATGAAGCAAGAGACTGAGACAGAGCCGCTGAGAGAA	4610	D	5631	GCTTACGAGAAACCTTATGTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	5690
Q	4558	TTTGAAGTGTGCTTCCCGAGCGGAGATGTATTTATGATGCGCTTGTGTGCTTCCGAA	4617	Q	5638	CAGGACGCTCTCAGAGAGAGAGCCCTGCGCGAGCGTACTGTGACATCCGGAACCGCGC	5697
D	4611	TTTGAAGTGTGCTTCCCGAGCGGAGATGTATTTATGATGCGCTTGTGTGCTTCCGAA	4670	D	5691	CAGGACGCTCTCAGAGAGAGAGCCCTGCGCGAGCGTACTGTGACATCCGGAACCGCGC	5750
Q	4618	CTGCAAAATACAGCAAGAGAGATGTCCCATACATCTGAAGATGAATCTCACCCGAC	4677	Q	5698	TACCTGGCCCTGCAATTTCTCTCAGAGAGGATTTACTTGGCGTCTCTATACAGAGATTA	5757
D	4671	CTGCAAAATACAGCAAGAGAGATGTCCCATACATCTGAAGATGAATCTCACCCGAC	4730	D	5751	TACCTGGCCCTGCAATTTCTCTCAGAGAGGATTTACTTGGCGTCTCTATACAGAGATTA	5810
Q	4678	ACCACTGTGGCCCGGAGAAACCTTACTTGTGTGCTCCAGCTTCCCTGACAAACAG	4737	Q	5758	TTAAGGCTATTGCTGTGCAAGAGAGAACTGTGAAAGAGTCCGACATGAAACACACCGG	5817
D	4731	ACCACTGTGGCCCGGAGAAACCTTACTTGTGTGCTCCAGCTTCCCTGACAAACAG	4790	D	5811	TTAAGGCTATTGCTGTGCAAGAGAGAACTGTGAAAGAGTCCGACATGAAACACACCGG	5870
Q	4738	CGCTGGGTCAACCGCTTGAATGATGTTGTGCAAGTGGAGAGATTTCTTAAAGAAAAAGCA	4797	Q	5818	GCGCGGTCCACCTCCGAGAGAGAGCCCAACAGAGAGGAGCCCAACAGTACAGAGAC	5877
D	4791	CGCTGGGTCAACCGCTTGAATGATGTTGTGCAAGTGGAGAGATTTCTTAAAGAAAAAGCA	4850	D	5871	GCGCGGTCCACCTCCGAGAGAGAGCCCAACAGAGAGGAGCCCAACAGTACAGAGAC	5930
Q	4798	GAAGCTGATGTAATGCTTGGAACTCCCTGTGAAACTGAGAGGTGATGACCTGCTA	4857	Q	5878	ATCACCAAGCGCTGTGCTTCAAGCCAGCGCCCGCAAGAGCCCAACAGCCCGCGAGAG	5937
D	4851	GAAGCTGATGTAATGCTTGGAACTCCCTGTGAAACTGAGAGGTGATGACCTGCTA	4910	D	5931	ATCACCAAGCGCTGTGCTTCAAGCCAGCGCCCGCAAGAGCCCAACAGCCCGCGAGAG	5990
Q	4858	GACATGAACCTGACAGCTGCCCTTCAAGTACCAAGTGTGTGTGTGTGTGTGTGTGT	4917	Q	5938	CCAAGCACACCCACCGCTTACCGGAGAGGAGGAGCGAGAGTGTGCGAGAGGACAAATGCTCT	5997
				D	5991	CCAAGCACACCCACCGCTTACCGGAGAGGAGGAGCGAGAGTGTGCGAGAGGACAAATGCTCT	6050

QY 5998 GGCGGCGGCTGGAGCGAGAGAGTCCCGCGCGGATCTACAGACCGCGAGAGAGCGG 6057
Db 6051 GGCGGCGGCTGGAGCGAGAGAGTCCCGCGCGGATCTACAGACCGCGAGAGAGCGG 6110
QY 6058 TCCCGCGGAGGCTGTTGAAGACAGACGAGGCGCGGCTGCTGGCGGAGCGGTGAGG 6117
Db 6111 TCCCGCGGAGGCTGTTGAAGACAGACGAGGCGCGGCTGCTGGCGGAGCGGTGAGG 6170
QY 6118 ACCCGCGTGTCCAGGTGAACAAGCTGTGGAGCAAGTCTTCAATATA 6165
Db 6171 ACCCGCGTGTCCAGGTGAACAAGCTGTGGAGCAAGTCTTCAATATA 6218

RESULT 5
US-09-964-956-10
Sequence 10, Application US/0964956
GENERAL INFORMATION:
APPLICANT: Gerlach, Valerie L
APPLICANT: MacDougall, John R
APPLICANT: Smtison, Glenda
APPLICANT: Miller, Isabelle
APPLICANT: Stone, David
APPLICANT: Gunther, Erik
APPLICANT: Ellerman, Karen
APPLICANT: Grosee, William M
APPLICANT: Alsodrook II, John P
APPLICANT: Lepley, Denise M
APPLICANT: Burgess, Catherine E
APPLICANT: Padigaru, Muralidhara
APPLICANT: Kekuda, Ramesh
APPLICANT: Spytek, Kimberly A
APPLICANT: Leach, Martin D
APPLICANT: Shimkets, Richard A
TITLE OF INVENTION: No. US20040043926A1 Proteins and Nucleic Acids Encoding Same
FILE REFERENCE: 21402-124
CURRENT APPLICATION NUMBER: US/09/964,956
CURRENT FILING DATE: 2001-09-26
PRIOR APPLICATION NUMBER: 60/235,631
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: 60/235,633
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: 60/235,808
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: 60/236,064
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: 60/236,065
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: 60/236,066
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: 60/236,135
PRIOR FILING DATE: 2000-09-28
PRIOR APPLICATION NUMBER: 60/237,434
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/238,321
PRIOR FILING DATE: 2000-10-05
PRIOR APPLICATION NUMBER: 60/238,399
PRIOR FILING DATE: 2000-10-06
PRIOR APPLICATION NUMBER: 60/238,396
PRIOR FILING DATE: 2000-10-06
PRIOR APPLICATION NUMBER: 60/276,667
PRIOR FILING DATE: 2001-03-16
PRIOR APPLICATION NUMBER: 60/294,823
PRIOR FILING DATE: 2001-05-31
PRIOR APPLICATION NUMBER: 60/304,868
NUMBER OF SEQ ID NOS: 127
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO: 10
LENGTH: 6189
TYPE: DNA
ORGANISM: Homo sapiens

US-09-964-956-10
Query Match 99.3%; Score 6121.8; DB 11; Length 6189;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 6145; Conservative 0; Mismatches 17; Indels 3; Gaps 1;
QY 1 ATGTTGAAGTTCAAAATATGAGAGCGGGAATCTTTGGATCTGCTGCTGAACCAATT 60
Db 1 ATGTTGAAGTTCAAAATATGAGAGCGGGAATCTTTGGATCTGCTGCTGAACCAATT 60
QY 61 GCCAGCGGAGCTCCAGGCTGAATCTGTTCTTCCAGGGAACCAACCTTTATGACTCAA 120
Db 61 GCCAGCGGAGCTCCAGGCTGAATCTGTTCTTCCAGGGAACCAACCTTTATGACTCAA 120
QY 121 CAGCAGATGCTCTCTTTCCGAGAGGATATAGATCCCTTTGTTCTTTGAA 180
Db 121 CAGCAGATGCTCTCTTTCCGAGAGGATATAGATCCCTTTGTTCTTTGAA 180
QY 181 GAATGAGTCAAGCTCTGCTGATGAAGATTAGCAGCTGAGCACTTTGTCGGAAAGTAT 240
Db 181 GAATGAGTCAAGCTCTGCTGATGAAGATTAGCAGCTGAGCACTTTGTCGGAAAGTAT 240
QY 241 TCCGACACATAGCTAGTTACAGAGCTCCAGCTTCCGCAAGAGCACTTCGAAGTCA 300
Db 241 TCCGACACATAGCTAGTTACAGAGCTCCAGCTTCCGCAAGAGCACTTCGAAGTCA 300
QY 301 AGTCTGTAGTGTGTGCTCACTTGTCTGAAGTCAAGTGTGAAGAGCAACCGG 360
Db 301 AGTCTGTAGTGTGTGCTCACTTGTCTGAAGTGTGAAGAGCAACCGG 360
QY 361 GACATCTATGCTATGAAGATGAAGAGAGCTTTATTTGGCCAGAGAGAGTTTCA 420
Db 361 GACATCTATGCTATGAAGATGAAGAGAGCTTTATTTGGCCAGAGAGAGTTTCA 420
QY 421 TTTTGGAGAGAGAGGAGCAATATATCCAGAGCAAGCCCGTGGATCCCGCAATTA 480
Db 421 TTTTGGAGAGAGAGGAGCAATATATCCAGAGCAAGCCCGTGGATCCCGCAATTA 480
QY 481 CAGTATGCTTTCAAGACAAATATCACTTTATCTGTGATGAATATAGCTGAGAGG 540
Db 481 CAGTATGCTTTCAAGACAAATATCACTTTATCTGTGATGAATATAGCTGAGAGG 540
QY 541 GACTTGTCTCACTTTTGAATATGATATGAGACAGATTAGTGAATACTGATACGTTT 600
Db 541 GACTTGTCTCACTTTTGAATATGATATGAGACAGATTAGTGAATACTGATACGTTT 600
QY 601 TACCTAGCTGAGCTATTTTGGCTGTTCACAGCTTCACTGATGAGATAGCTGATCA 660
Db 601 TACCTAGCTGAGCTATTTTGGCTGTTCACAGCTTCACTGATGAGATAGCTGATCA 660
QY 661 GACATCAAGCTGAGAACTTCTGTTGACGCAAGCAATCAATAGCTGTGATTTT 720
Db 661 GACATCAAGCTGAGAACTTCTGTTGACGCAAGCAATCAATAGCTGTGATTTT 720
QY 721 GGATCTGCGCGGAAATATGAAATCAAA--GATGAATGCAAACTCCGATTTGGACC 777
Db 721 GGATCTGCGCGGAAATATGAAATCAAA--GATGAATGCAAACTCCGATTTGGACC 777
QY 781 CCAAGTTACATAGCTCTCGAAGTGTGATGATGAACGGGAGATGAAAGGCACTTAC 840
Db 781 CCAAGTTACATAGCTCTCGAAGTGTGATGATGAACGGGAGATGAAAGGCACTTAC 840
QY 841 GGCTGGAAGTGAAGTGTGATGATGAGGAGGATTTGCTATGAGATATTTAGAGGAG 900
Db 841 GGCTGGAAGTGAAGTGTGATGATGAGGAGGATTTGCTATGAGATATTTAGAGGAG 900
QY 898 TCCCGCTTCCAGAGGAGAACTTCCGAGAACTTTCAATATTAATTAATTTCCAGGCG 957
Db 898 TCCCGCTTCCAGAGGAGAACTTCCGAGAACTTTCAATATTAATTAATTTCCAGGCG 957
QY 961 TTTTGAATTTCCAGATGACCCCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 1020
Db 961 TTTTGAATTTCCAGATGACCCCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 1020
QY 958 TTTTGAATTTCCAGATGACCCCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 1017
Db 958 TTTTGAATTTCCAGATGACCCCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 1017

QY 1021 TTGTTGCGGCGCAGAAAGAGAGACTGAAGTTTGAAGGCTTTGCTGCCATCTTTCTTC 1080
DB 1018 TTGTTGCGGCGCAGAAAGAGAGACTGAAGTTTGAAGGCTTTGCTGCCATCTTTCTTC 1077
QY 1081 TCTTAAATTTGATGGAACCAATCTGTAAGTCTCTCTCCCTCTGTTTCCACCTCTCAG 1140
DB 1078 TCTTAAATTTGATGGAACCAATCTGTAAGTCTCTCTCCCTCTGTTTCCACCTCTCAG 1137
QY 1141 TCTGACATGACACTCTCAATTTTGTATGAACACAGAGAAGATTGAGGTTTCACTCTCT 1200
DB 1138 TCTGACATGACACTCTCAATTTTGTATGAACACAGAGAAGATTGAGGTTTCACTCTCT 1197
QY 1201 CCGTGCCAGCTGAGCCCTCAGGCTTCTCGGCTGAAGAACTGCCGTTTGTGGGGTTTTCG 1260
DB 1198 CCGTGCCAGCTGAGCCCTCAGGCTTCTCGGCTGAAGAACTGCCGTTTGTGGGGTTTTCG 1257
QY 1261 TACAGCAGGCACTGAGGAGATTTCTTGTAAGTCTGTAAGTCTGTTGTGGGCTTGGACTCC 1320
DB 1258 TACAGCAGGCACTGAGGAGATTTCTTGTAAGTCTGTAAGTCTGTTGTGGGCTTGGACTCC 1317
QY 1321 CCTGCCAAGACTAGCTCTCAGTGAAGAAAGAACTTCTCATCAAAAGCAAAAGACTCAGAGAC 1380
DB 1318 CCTGCCAAGACTAGCTCTCAGTGAAGAAAGAACTTCTCATCAAAAGCAAAAGACTCAGAGAC 1377
QY 1381 TCTCAGGACAAGTGTCACAAGATGAGCAAGAAATGACCCGGTTACATCGAGAGTGTCA 1440
DB 1378 TCTCAGGACAAGTGTCACAAGATGAGCAAGAAATGACCCGGTTACATCGAGAGTGTCA 1437
QY 1441 GAGGTGAGGCTGTGCTTGTAGTCAAGAGAGAGTGAAGCTGAAGGCTCTGAGACTCAGAGA 1500
DB 1438 GAGGTGAGGCTGTGCTTGTAGTCAAGAGAGAGTGAAGGCTCTGAGACTCAGAGA 1497
QY 1501 TCCCTCCTGAGAGAGGACTTGTACTACATCAACAATGCAATAGCTTAAAGCGAAGT 1560
DB 1498 TCCCTCCTGAGAGAGGACTTGTACTACATCAACAATGCAATAGCTTAAAGCGAAGT 1557
QY 1561 TTGAGCAGACACGAGATGAGAGTGTCCAGAGAGATGACAAAGCACTGACGCTTCTCAT 1620
DB 1558 TTGAGCAGACACGAGATGAGAGTGTCCAGAGAGATGACAAAGCACTGACGCTTCTCAT 1617
QY 1621 GATATCAGAGACAGAGCCCGAAGCTCCAAAGAAATCAAGAGCAGAGTACAGAGCTCAA 1680
DB 1618 GATATCAGAGACAGAGCCCGAAGCTCCAAAGAAATCAAGAGCAGAGTACAGAGCTCAA 1677
QY 1681 GTGGAAGAAATGAGTGTGATGATGAATCAGTTGGAAGAGATCTTGTCTCAGCAAGAGA 1740
DB 1678 GTGGAAGAAATGAGTGTGATGATGAATCAGTTGGAAGAGATCTTGTCTCAGCAAGAGA 1737
QY 1741 CCGAGTGTCTCTACGAATCTGAGACTGAGAGAGTCTCGGCTTGTGCTGAAGAAATTCAG 1800
DB 1738 CCGAGTGTCTCTACGAATCTGAGACTGAGAGAGTCTCGGCTTGTGCTGAAGAAATTCAG 1797
QY 1801 CGGAAGCGACAGAAATGTCAGCATTAACCTGTTGAAGCTTAAGATCAAGGGAGCTGAA 1860
DB 1798 CGGAAGCGACAGAAATGTCAGCATTAACCTGTTGAAGCTTAAGATCAAGGGAGCTGAA 1857
QY 1861 GTGGAAGAAATGTCAGAACTGGAAGAAATCAATGCTGAGCAGAGCTCAAAATTCAGAG 1920
DB 1858 GTGGAAGAAATGTCAGAACTGGAAGAAATCAATGCTGAGCAGAGCTCAAAATTCAGAG 1917
QY 1921 CTCGAAGAGAACTGGAAGAGGCTGTAAAGCCAGCAGAGGCGACCGAGCTCTCAG 1980
DB 1918 CTCGAAGAGAACTGGAAGAGGCTGTAAAGCCAGCAGAGGCGACCGAGCTCTCAG 1977
QY 1981 AATATCCGCGACGAAGAGAGCGAGCGAGAGGAGCTGAGAAAGCTGCAAGACCGAGAG 2040
DB 1978 AATATCCGCGACGAAGAGAGCGAGCGAGAGGAGCTGAGAAAGCTGCAAGACCGAGAG 2037
QY 2041 GATTTCTTGAAGGCACTCAAGAAAGTGTGTGAAGCTGAAGAACCCCGCATTTCTCTG 2100
DB 2038 GATTTCTTGAAGGCACTCAAGAAAGTGTGTGAAGCTGAAGAACCCCGCATTTCTCTG 2097

QY 2101 GAGAACAAAGTAAGAGACTAGAGACCAATGAGCGTGAAGAAAACAGACTGAAGATGAC 2160
DB 2098 GAGAACAAAGTAAGAGACTAGAGACCAATGAGCGTGAAGAAAACAGACTGAAGATGAC 2157
QY 2161 ATCCAGCAAAATCCCAACAGATCCAGCAGATGGCTGATATAAATTTGAGAGCTGAAGAG 2220
DB 2158 ATCCAGCAAAATCCCAACAGATCCAGCAGATGGCTGATATAAATTTGAGAGCTGAAGAG 2217
QY 2221 AAACATCGGAGGCGCCAAAGTCTGAGCCAGCAGCTAGAAAGTGCACCTGAAAACAGAAAG 2280
DB 2218 AAACATCGGAGGCGCCAAAGTCTGAGCCAGCAGCTAGAAAGTGCACCTGAAAACAGAAAG 2277
QY 2281 CAGCAGTATGAGAAAAGATTAAAGTGTGAACAATCAGATTAAGAAAGACCTGGCTGAC 2340
DB 2278 CAGCAGTATGAGAAAAGATTAAAGTGTGAACAATCAGATTAAGAAAGACCTGGCTGAC 2337
QY 2341 AAGGAGCACTGAGAGAACATGATGAGAGACAAGAGAGAGGCGCCATGAGAAAGGCAAA 2400
DB 2338 AAGGAGCACTGAGAGAACATGATGAGAGACAAGAGAGAGGCGCCATGAGAAAGGCAAA 2397
QY 2401 ATTCTCAGCGAAGAGAGCGATGATCAATGCTATGATTCAGATCAGATCCCTGAAA 2460
DB 2398 ATTCTCAGCGAAGAGAGCGATGATCAATGCTATGATTCAGATCAGATCCCTGAAA 2457
QY 2461 CAGAGATTTGGAATCTGTTGAAGCCAAATTAACCTGACCAATATAGCAGCTTTTAAAC 2520
DB 2458 CAGAGATTTGGAATCTGTTGAAGCCAAATTAACCTGACCAATATAGCAGCTTTTAAAC 2517
QY 2521 CAAAGGAACATGAAGGCGCAAGAAAGATGATTTCTGAACTCAGAGGCAACGAATTTTAC 2580
DB 2518 CAAAGGAACATGAAGGCGCAAGAAAGATGATTTCTGAACTCAGAGGCAACGAATTTTAC 2577
QY 2581 CTGAGACACAGGCTGGAGAGTGTGAAGGCCAGAAACGAAATCTGAGAGCAGCTGAG 2640
DB 2578 CTGAGACACAGGCTGGAGAGTGTGAAGGCCAGAAACGAAATCTGAGAGCAGCTGAG 2637
QY 2641 AAGATCAGCCACCAAGACCAAGTGAACAAGATGCGTCTGGAATCTGAGACAAAGATTG 2700
DB 2638 AAGATCAGCCACCAAGACCAAGTGAACAAGATGCGTCTGGAATCTGAGACAAAGATTG 2697
QY 2701 CCGGAGGTCAAGTCTAGAGCAAGAGAGCAAAATCTGAAGCTCAAGCGCCAGCTCAAGAG 2760
DB 2698 CCGGAGGTCAAGTCTAGAGCAAGAGAGCAAAATCTGAAGCTCAAGCGCCAGCTCAAGAG 2757
QY 2761 CTACAGCTCTCCCTGCGAGAGCGGAGTCAAGTTGACAGCCCTGACAGGCTGCAGGGCG 2820
DB 2758 CTACAGCTCTCCCTGCGAGAGCGGAGTCAAGTTGACAGCCCTGACAGGCTGCAGGGCG 2817
QY 2821 GCCCTGAGAGCGAGCTTGCAGAGCGAAGACAGAGCTGGAAGAGACCAAGCAAGAGCT 2880
DB 2818 GCCCTGAGAGCGAGCTTGCAGAGCGAAGACAGAGCTGGAAGAGACCAAGCAAGAGCT 2877
QY 2881 GAAGAGAGATCCAGGCACTCAAGGCAATAGAGATGAATCCAGCGCCAAATTTGATGCT 2940
DB 2878 GAAGAGAGATCCAGGCACTCAAGGCAATAGAGATGAATCCAGCGCCAAATTTGATGCT 2937
QY 2941 CTTGCTAAGAGCTGTACTGTATATCAAGACCTGGAAGAGCGAGCTTAAACAGCTGACCGAG 3000
DB 2938 CTTGCTAAGAGCTGTACTGTATATCAAGACCTGGAAGAGCGAGCTTAAACAGCTGACCGAG 2997
QY 3001 GACACGCTGAATCTCAACCAACCAAACTTCTACTTGTCAAAACAATCGATGAGGCTTCT 3060
DB 2998 GACACGCTGAATCTCAACCAACCAAACTTCTACTTGTCAAAACAATCGATGAGGCTTCT 3057
QY 3061 GGGCGCAAGCAAGAGATTGTACAACTGCGAAGTGAATGGAATCTTCCGCGGGAGATC 3120
DB 3058 GGGCGCAAGCAAGAGATTGTACAACTGCGAAGTGAATGGAATCTTCCGCGGGAGATC 3117
QY 3121 ACCGAAGAGAGATGAGCTTACAGCGAGCAAGCAAGCAATGAGGCTTGAAGACCAAG 3180
DB 3118 ACCGAAGAGAGATGAGCTTACAGCGAGCAAGCAAGCAATGAGGCTTGAAGACCAAG 3177
QY 3181 TGCAACATGCTGAGAGAACAGATCATGATTTGAGAGCCCTTAAACGATGAGCTGAGAA 3240

Db 3178 TGCACCATGCTGGAGGAAACAGTCAATGATTTGAGAGCCCTTAACGATGAGCTTGA 3237
Qy 3241 AAGAGCGGCACTGGAGAGCCCTGAGAGCCCTCTGGGTGATGAGAAATCCCACTTTGAG 3300
Db 3238 AAGAGCGGCACTGGAGAGCCCTGAGAGCCCTCTGGGTGATGAGAAATCCCACTTTGAG 3297
Qy 3301 TGTGGGTTTGGAGAGCTGGAGAGATGCTGAGACCCGAGAAACAGACGAGGCGAGGCC 3360
Db 3298 TGTGGGTTTGGAGAGCTGGAGAGATGCTGAGACCCGAGAAACAGACGAGGCGAGGCC 3357
Qy 3361 GATCAGCGATACCCAGATCTGCCAGATGGTGGAGCTGGCAGTGAAGAGACAAAGGCT 3420
Db 3358 GATCAGCGATACCCAGATCTGCCAGATGGTGGAGCTGGCAGTGAAGAGACAAAGGCT 3417
Qy 3421 GAGATTCTGCTCTGACAGAGCTCTCAAAAGACGAGAGCTGAAGCCGAGAGCTCTCT 3480
Db 3418 GAGATTCTGCTCTGACAGAGCTCTCAAAAGACGAGAGCTGAAGCCGAGAGCTCTCT 3477
Qy 3481 GACAAGCTCAATGACCTGGAGAGAAAGATGCTATGCTTTGAATGAATGCCGAACTTA 3540
Db 3478 GACAAGCTCAATGACCTGGAGAGAAAGATGCTATGCTTTGAATGAATGCCGAACTTA 3537
Qy 3541 CACGAGAGCTGGAGACTGAACGAGAGCTCAAAACAGAGGCTTCTGAAAGCAAGCCAA 3600
Db 3538 CACGAGAGCTGGAGACTGAACGAGAGCTCAAAACAGAGGCTTCTGAAAGCAAGCCAA 3597
Qy 3601 TTAACAGACGATGGAACCTGCAAAAAATCACATTTCCGCTGACTCAAGACTGCA 3660
Db 3598 TTAACAGACGATGGAACCTGCAAAAAATCACATTTCCGCTGACTCAAGACTGCA 3657
Qy 3661 GAAGCTCTGATGCGGCTGATCTACTGAGACAGAAAGAGTGAAGTGAAGTCAAGCTG 3720
Db 3658 GAAGCTCTGATGCGGCTGATCTACTGAGACAGAAAGAGTGAAGTGAAGTCAAGCTG 3717
Qy 3721 GAAAAATTCAAGGTTCTCTATTCTCATGAAAAAGTGAAGGCAATATTTCTCA 3780
Db 3718 GAAAAATTCAAGGTTCTCTATTCTCATGAAAAAGTGAAGGCAATATTTCTCA 3777
Qy 3781 CAAACCAAACTCATTTATTTCTGCAAGCCAAATGAGCCACTGTAAAAAGAAAAAG 3840
Db 3778 CAAACCAAACTCATTTATTTCTGCAAGCCAAATGAGCCACTGTAAAAAGAAAAAG 3837
Qy 3841 GTTCCCTCAGTACATGAGCTGAAGCTGACCTTGAAGAGAAAGTCCCTGCTGCA 3900
Db 3838 GTTCCCTCAGTACATGAGCTGAAGCTGACCTTGAAGAGAAAGTCCCTGCTGCA 3897
Qy 3901 GAGCTAGAGAAAGCCCTTCAGAAAGCCGATGAGCTCCGCTCCGCGGAGAGAGCT 3960
Db 3898 GAGCTAGAGAAAGCCCTTCAGAAAGCCGATGAGCTCCGCTCCGCGGAGAGAGCT 3957
Qy 3961 GCCCAGCGAAAGCAACGAGCAACCAACCATTCACGCGCAGCCAGCGAGGAGCAG 4020
Db 3958 GCCCAGCGAAAGCAACGAGCAACCAACCATTCACGCGCAGCCAGCGAGGAGCAG 4017
Qy 4021 ATGCGCATGTCGCGCATGCTGGGTGCGAGAGCAACGAGCCAGTGCATGAGCTGCTG 4080
Db 4018 ATGCGCATGTCGCGCATGCTGGGTGCGAGAGCAACGAGCCAGTGCATGAGCTGCTG 4077
Qy 4081 GCCCGCCCATTCAGCGCGAGAAAGAGTCTTCAATCCAGAGAAATTTAGTGGCGCTTT 4140
Db 4078 GCCCGCCCATTCAGCGCGAGAAAGAGTCTTCAATCCAGAGAAATTTAGTGGCGCTTT 4137
Qy 4141 AAGGAACGATCACCAATATTTCTCAACGATTCAGGATGAGACTGAACATGCGAGCC 4200
Db 4138 AAGGAACGATCACCAATATTTCTCAACGATTCAGGATGAGACTGAACATGCGAGCC 4197
Qy 4201 ACAAAGTGTGCTGTGTGTGATACGCTGCACTTTGAGCGCCAGGATCCAAATGTCTC 4260
Db 4198 ACAAAGTGTGCTGTGTGTGATACGCTGCACTTTGAGCGCCAGGATCCAAATGTCTC 4257
Qy 4261 GAAATGTCAGGTGATGTCACCCCAAGTCTCAAGTGTTCAGCCCACTGCGGCTTG 4320

Db 4258 GAAATGTCAGGTGATGTCACCCCAAGTCTCCAGCTGCTTCCAGCCACTGCGGCTTG 4317
Qy 4321 CTTGCTGAATATGCAACAATTCAACGAGGCTTCTGCGGTGACAAATATGAATCTCCCA 4380
Db 4318 CTTGCTGAATATGCAACAATTCAACGAGGCTTCTGCGGTGACAAATATGAATCTCCCA 4377
Qy 4381 GGTCTCCAGACCAAGAGCCCAAGAGCTTGTGACCTTGAAGGGTGGATGAAGTCCCC 4440
Db 4378 GGTCTCCAGACCAAGAGCCCAAGAGCTTGTGACCTTGAAGGGTGGATGAAGTCCCC 4437
Qy 4441 AGGAATTAACAAAGAGACGAAAGCTGGGACAGGAAGTACAATTCCTGGAGGGATCA 4500
Db 4438 AGGAATTAACAAAGAGACGAAAGCTGGGACAGGAAGTACAATTCCTGGAGGGATCA 4497
Qy 4501 AAGTCTCATTTATGATCAATGAAGCCAGAGAGCTGACAGAGGCGGTGAGAAATTT 4560
Db 4498 AAGTCTCATTTATGATCAATGAAGCCAGAGAGCTGACAGAGGCGGTGAGAAATTT 4557
Qy 4561 GAGCTGTGCTTCCCAACGAGGATGTAATTCATGCTGCGGTGGTCTTCCGAACTC 4620
Db 4558 GAGCTGTGCTTCCCAACGAGGATGTAATTCATGCTGCGGTGGTCTTCCGAACTC 4617
Qy 4621 GCAATTAACGCGCAAGACAGATGTCCTCATATCTGAAGTGAATTCACCCGCAACC 4680
Db 4618 GCAATTAACGCGCAAGACAGATGTCCTCATATCTGAAGTGAATTCACCCGCAACC 4677
Qy 4681 ACCTGTGACCGCGGAGAAACCTCTACTGCTAGCTCCAGCTTCCCTGACAAACAGCGC 4740
Db 4678 ACCTGTGACCGCGGAGAAACCTCTACTGCTAGCTCCAGCTTCCCTGACAAACAGCGC 4737
Qy 4741 TGGGTACCGGCTTGAATCAAGTTGTGCGAGGTGGAGAGTTCTTGGGAAAAAGCAGAA 4800
Db 4738 TGGGTACCGGCTTGAATCAAGTTGTGCGAGGTGGAGAGTTCTTGGGAAAAAGCAGAA 4797
Qy 4801 GCTGATGCTAAACTGCTTGAAGCTCCGCTGAAACCTGGAAGGTGATGACCGTCTAGAC 4860
Db 4798 GCTGATGCTAAACTGCTTGAAGCTCCGCTGAAACCTGGAAGGTGATGACCGTCTAGAC 4857
Qy 4861 ATGAACTGACGCTGCTTCAAGTGAACAGGTGCTGCTGGGCAACGAGAGAGGCTC 4920
Db 4858 ATGAACTGACGCTGCTTCAAGTGAACAGGTGCTGCTGGGCAACGAGAGAGGCTC 4917
Qy 4921 TACGCTGTAATGTCTTGAAGAACTCCCTAACCCATGTCCTCCAGAAATGAGAGCTTTC 4980
Db 4918 TACGCTGTAATGTCTTGAAGAACTCCCTAACCCATGTCCTCCAGAAATGAGAGCTTTC 4977
Qy 4981 CAAATTTATATATCAAGGACCTGAGAAAGTACTATGATGCGAGAGAGAGGCGGCA 5040
Db 4978 CAAATTTATATATCAAGGACCTGAGAAAGTACTATGATGCGAGAGAGAGGCGGCA 5037
Qy 5041 CTGTGCTTGTGAGAGTGAAGAAAGTGAACAGTCCCTGAGCCAGTCCCACTGCTGCGC 5100
Db 5038 CTGTGCTTGTGAGAGTGAAGAAAGTGAACAGTCCCTGAGCCAGTCCCACTGCTGCGC 5097
Qy 5101 CAGCCCGACATTCACCCAACTTTTGAAGCTGTCAAGGCTGCCACTTGTTTGGGCGCA 5160
Db 5098 CAGCCCGACATTCACCCAACTTTTGAAGCTGTCAAGGCTGCCACTTGTTTGGGCGCA 5157
Qy 5161 GGCAGAGTTGAGAAAGGCTGCGCATGCTGAGAGCAAGCCAGAAAGTCCCTCAATTCG 5220
Db 5158 GGCAGAGTTGAGAAAGGCTGCGCATGCTGAGAGCAAGCCAGAAAGTCCCTCAATTCG 5217
Qy 5221 CGCTACAAAGAAACCTGACGAAATATGCAATCCGAGAAAGAGATAGAGCTCAGAGCC 5280
Db 5218 CGCTACAAAGAAACCTGACGAAATATGCAATCCGAGAAAGAGATAGAGCTCAGAGCC 5277
Qy 5281 TGCAGCTGTATTCACCTTCAACCAATTTACAGTATCTCAATTGGAACCAATTAATTTACGAA 5340
Db 5278 TGCAGCTGTATTCACCTTCAACCAATTTACAGTATCTCAATTGGAACCAATTAATTTACGAA 5337
Qy 5341 ATGCAATGAAGCAATACAGCTGAGAGAAATTCCTGATTAAGATGACCAATTCCTTGGCA 5400
Db 5338 ATGCAATGAAGCAATACAGCTGAGAGAAATTCCTGATTAAGATGACCAATTCCTTGGCA 5397

```

QY 5401 CCGTGTGTTTGGCCGCTCTTCCAAAGCTTCCCTGTCTCAATGTCAGGTGAACAGC 5460
DB 5398 CCGTGTGTTTGGCCGCTCTTCCAAAGCTTCCCTGTCTCAATGTCAGGTGAACAGC 5457
QY 5461 GCAGGGGAGGAGGAGGAGTCTGTGTGTTCACCAATTTGAGAGTGTGTGTGTGTCT 5520
DB 5458 GCAGGGGAGGAGGAGGAGTCTGTGTGTTCACCAATTTGAGAGTGTGTGTGTGTCT 5517
QY 5521 TACGGAAGAGTACGCGCAGACGATCTCAAGTGAAGTGTGTGTGTGTGTGTGTGTGT 5580
DB 5518 TACGGAAGAGTACGCGCAGACGATCTCAAGTGAAGTGTGTGTGTGTGTGTGTGTGT 5577
QY 5581 TACGGAAGAGTACGCGCAGACGATCTCAAGTGAAGTGTGTGTGTGTGTGTGTGTGT 5640
DB 5578 TACGGAAGAGTACGCGCAGACGATCTCAAGTGAAGTGTGTGTGTGTGTGTGTGTGT 5637
QY 5641 GCAGGCTCTTCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 5700
DB 5638 GCAGGCTCTTCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 5697
QY 5701 CCGGCGCCCTGCAATTTCTCTCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 5760
DB 5698 CCGGCGCCCTGCAATTTCTCTCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 5757
QY 5761 AGGCTCATTTGCTGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 5820
DB 5758 AGGCTCATTTGCTGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 5817
QY 5821 CCGTCCACCTTCCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 5880
DB 5818 CCGTCCACCTTCCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 5877
QY 5881 ACCAAGGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 5940
DB 5878 ACCAAGGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 5937
QY 5941 AGCAGACCCCAACCGCTACCGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 6000
DB 5938 AGCAGACCCCAACCGCTACCGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 5997
QY 6001 CGCCCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 6060
DB 5998 CGCCCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 6057
QY 6061 CCGCGGAGGCTGTTTGAAGAGAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 6120
DB 6058 CCGCGGAGGCTGTTTGAAGAGAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 6117
QY 6121 CCGGCTGTCCGAGGTGAACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 6165
DB 6118 CCGGCTGTCCGAGGTGAACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 6162

```

RESULT 6
US-10-262-511-1
Sequence 1, Application US/10262511
Publication No. US20040038223A1

GENERAL INFORMATION:

APPLICANT: Smithson, Glenda
APPLICANT: Miller, Isabelle
APPLICANT: Peyman, John A.
APPLICANT: Kemeah
APPLICANT: Ju, Jinfang
APPLICANT: Li, Li
APPLICANT: Guo, Xiaojia (Saeha)
APPLICANT: Paturajan, Meera
APPLICANT: Spytek, Kimberly A.
APPLICANT: Edinger, Shlomit R.
APPLICANT: Ellerman, Karen
APPLICANT: Malyankar, Uriel M.
APPLICANT: Ort, Tatiana
APPLICANT: Gorman, Linda

```

APPLICANT: Zerhusen, Bryan D.
APPLICANT: Anderson, David W.
APPLICANT: Zhong, Mei
APPLICANT: Catterton, Elina
APPLICANT: Ji, Weizhen
APPLICANT: Miller, Charles E.
APPLICANT: Rastelli, Luca
APPLICANT: Stone, David J.
APPLICANT: Pena, Carol E. A.
APPLICANT: Shenoy, Suresh G.
APPLICANT: Shinkets, Richard A.
APPLICANT: Rothenberg, Mark E.
APPLICANT: Leach, Martin D.
APPLICANT: Agee, Michele L.
APPLICANT: Berghs, Constance
TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
FILE REFERENCE: 21402-462C
CURRENT APPLICATION NUMBER: US/10/262,511
CURRENT FILING DATE: 2003-05-28
PRIOR APPLICATION NUMBER: 60/326,483
PRIOR FILING DATE: 2001-10-02
PRIOR APPLICATION NUMBER: 60/373,815
PRIOR FILING DATE: 2002-04-19
PRIOR APPLICATION NUMBER: 60/327,917
PRIOR FILING DATE: 2001-10-09
PRIOR APPLICATION NUMBER: 60/381,642
PRIOR FILING DATE: 2002-05-17
PRIOR APPLICATION NUMBER: 60/328,029
PRIOR FILING DATE: 2002-10-09
PRIOR APPLICATION NUMBER: 60/381,038
PRIOR FILING DATE: 2002-05-16
PRIOR APPLICATION NUMBER: 60/328,056
PRIOR FILING DATE: 2001-10-09
PRIOR APPLICATION NUMBER: 60/373,260
PRIOR FILING DATE: 2002-04-17
PRIOR APPLICATION NUMBER: 60/373,826
PRIOR FILING DATE: 2002-04-19
PRIOR APPLICATION NUMBER: 60/327,435
PRIOR FILING DATE: 2001-10-05
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 439
SOFTWARE: CuroSeqList version 0.1
SEQ ID NO 1
LENGTH: 6189
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (1)..(6159)
US-10-262-511-1

```

Query Match 99.3%; Score 6121.8; DB 17; Length 6189;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 6145; Conservative 0; Mismatches 17; Indels 3; Gaps 1;

```

QY 1 ATGTTGAAGTTCAATATGAGCGCGGAGATCTTTGATGTCGTGCTGAACCAATT 60
DB 1 ATGTTGAAGTTCAATATGAGCGCGGAGATCTTTGATGTCGTGCTGAACCAATT 60
QY 61 GCCAGCGGCGCTCCAGGCTGAATCTTTCTTCCAGGGGAAACCACTTTATGACTCA 120
DB 61 GCCAGCGGCGCTCCAGGCTGAATCTTTCTTCCAGGGGAAACCACTTTATGACTCA 120
QY 121 CAGGAGATGTCCTCTCTTTCCCGGAGAGGAGATATTAGTGCCTTTGTTCTTTGAA 180
DB 121 CAGGAGATGTCCTCTCTTTCCCGGAGAGGAGATATTAGTGCCTTTGTTCTTTGAA 180
QY 181 GAATGACAGTACGCTGCTGATGAGATTAAAGCAGTGAACATTTGTCCGAAAGTAT 240
DB 181 GAATGACAGTACGCTGCTGATGAGATTAAAGCAGTGAACATTTGTCCGAAAGTAT 240
QY 241 TCGACACCATAGCTGAGTTACAGAGAGCTCCAGCTTCCGCAAGAGACTTGAAGTCA 300
DB 241 TCGACACCATAGCTGAGTTACAGAGAGCTCCAGCTTCCGCAAGAGACTTGAAGTCA 300

```


Db 241 TCCGACACCATAGCTGAGTTACAGGAGCTCCAGCCTTCGCGAAGAGACTTCGAAATCAGA 300
QY 301 AGTCTTGAGAGTTGGTCACTTTTGCTGAAAGTGAGTGATGAAGAGAAAGCAACCGGG 360
Db 301 AGTCTTGAGAGTTGGTCACTTTTGCTGAAAGTGAGTGATGAAGAGAAAGCAACCGGG 360
QY 361 GACATCTATGCTATGAAGATGATGAAGAGAGGCTTTATTTGGCCCGAGAGCAAGTTTCA 420
Db 361 GACATCTATGCTATGAAGATGATGAAGAGAGGCTTTATTTGGCCCGAGAGCAAGTTTCA 420
QY 421 TTTTGTGAGGAAGAGCGGAACATATATCTCGAAGCACAAGCCGCTGATCCCAATTA 480
Db 421 TTTTGTGAGGAAGAGCGGAACATATATCTCGAAGCACAAGCCGCTGATCCCAATTA 480
QY 481 CAGATACCTTTTCAGGACAAAATACACCTTTATCTGGTGAATGAAATTCAGCTGAGGG 540
Db 481 CAGATACCTTTTCAGGACAAAATACACCTTTATCTGGTGAATGAAATTCAGCTGAGGG 540
QY 541 GACTGCTGTCATCTTTGAATAGATAGAGGACAGTTAGATGAATAACCTGATCACTTT 600
Db 541 GACTGCTGTCATCTTTGAATAGATAGAGGACAGTTAGATGAATAACCTGATCACTTT 600
QY 601 TACCTAGCTGAGTGAATTTTGCTGTTCACAGCCTTCATGATGGGATACGTGCATCGA 660
Db 601 TACCTAGCTGAGTGAATTTTGCTGTTCACAGCCTTCATGATGGGATACGTGCATCGG 660
QY 661 GACATCAAGCCTGAGAACTTCTGTTGACCGGACAGAGACATCAAGCTGTGTGATTTT 720
Db 661 GACATCAAGCCTGAGAACTTCTGTTGACCGGACAGAGACATCAAGCTGTGTGATTTT 720
QY 721 GGATCTGCGCGGAAAATGAATTCAAAACAGATGGTGAAATGCCAACTCCGATTGGGACC 780
Db 721 GGATCTGCGCGGAAAATGAATTCAAAACA--GGTGATGCCAACTCCGATTGGGACC 777
QY 781 CCGAATTAACATGAGCTCCTGAAAGTCTGACTGTGATGAAACGGAGATGAAAAGGCACTAC 840
Db 778 CCGAATTAACATGAGCTCCTGAAAGTCTGACTGTGATGAAACGGAGATGAAAAGGCACTAC 837
QY 841 GGGCTGAGCTGTACTGTGTGTGATGAGGCGGTGATTTGCTTAATGATGATTTATGGAGA 900
Db 838 GGGCTGAGCTGTGACTGTGTGTGATGAGGCGGTGATTTGCTTAATGATGATTTATGGAGA 897
QY 901 TCCCCCTCGCAGAGGGAACCTCTGCAGAACTTCAATTAACATTAATGAATTTCCAGCGG 960
Db 898 TCCCCCTCGCAGAGGGAACCTCTGCAGAACTTCAATTAACATTAATGAATTTCCAGCGG 957
QY 961 TTTTGGAAATTTCCAGATGACCCCAAGTGAAGAGTGAAGTCTTTGATCTGATTCAGAG 1020
Db 958 TTTTGGAAATTTCCAGATGACCCCAAGTGAAGAGTGAAGTCTTTGATCTGATTCAGAG 1017
QY 1021 TTGTTGTGCGGCGAAGAAAGAGACTGAAGTGAAGTCTTTGCTGCCATCCTTCTTC 1080
Db 1018 TTGTTGTGCGGCGAAGAAAGAGACTGAAGTGAAGTCTTTGCTGCCATCCTTCTTC 1077
QY 1081 TCTAAAAATTGACTGGAACAACATTGCTAACTCTCCCTCCCTCGTTCCTCCACCTCAAG 1140
Db 1078 TCTAAAAATTGACTGGAACAACATTGCTAACTCTCCCTCCCTCGTTCCTCCACCTCAAG 1137
QY 1141 TCTGACGATGACCTTCAATTTTGATGAACAGAGAAATTTGAGGTTTCATCTCT 1200
Db 1138 TCTGACGATGACCTTCAATTTTGATGAACAGAGAAATTTGAGGTTTCATCTCTCT 1197
QY 1201 CCGTGCAGCTGAGCCCTCAGGCTTCTCGGGTGAAGAACTGCCGTTTGTGGGGTTTTCG 1260
Db 1198 CCGTGCAGCTGAGCCCTCAGGCTTCTCGGGTGAAGAACTGCCGTTTGTGGGGTTTTCG 1257
QY 1261 TACAGCAAGGCACTGGGGATTTCTTGTTAATCTGAGTCTGTTGTGTCTGGGCTTCC 1320
Db 1258 TACAGCAAGGCACTGGGGATTTCTTGTTAATCTGAGTCTGTTGTGTCTGGGCTTCC 1317
QY 1321 CCTGCCAAGACTAGCTCAATGAAAAAAATTTCTCATCAAAAGCAAGAGCTACAGAC 1380
Db 1318 CCTGCCAAGACTAGCTCAATGAAAAAAATTTCTCATCAAAAGCAAGAGCTACAGAC 1377

QY 1381 TCTCAGGACAAGTGTCACAAGATGAGACAGAAAATGACCCGGTTTACATCGAGAGTGTCA 1440
Db 1378 TCTCAGGACAAGTGTCACAAGATGAGACAGAAAATGACCCGGTTTACATCGAGAGTGTCA 1437
QY 1441 GAGGTGAGGCTGTGCTTACTCAGAAAGAGTGAAGCTTGAAGCTTCTGAGCTCAGAGA 1500
Db 1438 GAGGTGAGGCTGTGCTTACTCAGAAAGAGTGAAGCTTGAAGCTTCTGAGACTCAGAGA 1497
QY 1501 TCCCTCTGAGAGAGGACTTGTGCTTACATCATCAAGAAATGCAAGTCTTAAAGCGAAGT 1560
Db 1498 TCCCTCTGAGAGAGGACTTGTGCTTACTCAGATCAAGAAATGCAAGTCTTAAAGCGAAGT 1557
QY 1561 TTGAGCAAGCAACGATGAGGCTGTCCAGAGAGATGACAAAGCACTGCAGTCTTCCAT 1620
Db 1558 TTGAGCAAGCAACGATGAGGCTGTCCAGAGAGATGACAAAGCACTGCAGTCTTCCAT 1617
QY 1621 GATATCAGAGAGCAGAGCCGGAAGCTTCAGAAATCAAGAGCAGAGTACAGGCTCAA 1680
Db 1618 GATATCAGAGAGCAGAGCCGGAAGCTTCAGAAATCAAGAGCAGAGTACAGGCTCAA 1677
QY 1681 GTGGAAGAAATGAGTTGATGATGAATCAAGTGAAGAGGATCTTGCTCAGCAAGAGA 1740
Db 1678 GTGGAAGAAATGAGTTGATGATGAATCAAGTGAAGAGGATCTTGCTCAGCAAGAGA 1737
QY 1741 CCGAGTGAATCTTACGAATCTGAGCTGAGAGTCTCGGCTTGCTGTAAGAAATTCAG 1800
Db 1738 CCGAGTGAATCTTACGAATCTGAGCTGAGAGTCTCGGCTTGCTGTAAGAAATTCAG 1797
QY 1801 CCGAAGCCGACAGAAATGTCAAGCTAAATCTGTTGAAGGCTTAAGATCAAGGAAAGCTGAA 1860
Db 1798 CCGAAGCCGACAGAAATGTCAAGCTAAATCTGTTGAAGGCTTAAGATCAAGGAAAGCTGAA 1857
QY 1861 GTGGAGAAATATGCGGAACTGGAAGAAATCAATGTCTGACAGAGCTCAAAATTCAGAG 1920
Db 1858 GTGGAGAAATATGCGGAACTGGAAGAAATCAATGTCTGACAGAGCTCAAAATTCAGAG 1917
QY 1921 CTCCAAGAGAAATCTGAGAGAGGCTGTAAAGCCAGACCGAGGCCACCGAGCTGTCCAG 1980
Db 1918 CTCCAAGAGAAATCTGAGAGAGGCTGTAAAGCCAGACCGAGGCCACCGAGCTGTCCAG 1977
QY 1981 AATATCCGCAAGCAAGAGAGAGGAGCTGGAAGAGCTGGAAGAGCTGCAAGCCAGAG 2040
Db 1978 AATATCCGCAAGCAAGAGAGAGGAGCTGGAAGAGCTGGAAGAGCTGCAAGCCAGAG 2037
QY 2041 GATTCTTGAAGGCAATCGAAGAAAGCTGTGGAAGCTGAGGAACGCCCAATCTCTG 2100
Db 2038 GATTCTTGAAGGCAATCGAAGAAAGCTGTGGAAGCTGAGGAACGCCCAATCTCTCTG 2097
QY 2101 GAGAACAAAGTAAAGAGCTAGAGACCATGAGCGTGAAGAAAACAGACTGAGAGATGAC 2160
Db 2098 GAGAACAAAGTAAAGAGCTAGAGACCATGAGCGTGAAGAAAACAGACTGAGAGATGAC 2157
QY 2161 ATCCAGACAAATCCCAACAGATCCAGAGATGCTGATTAATTTCTGAGGCTCGAAGAG 2220
Db 2158 ATCCAGACAAATCCCAACAGATCCAGAGATGCTGATTAATTTCTGAGGCTCGAAGAG 2217
QY 2221 AAACATCGGAGGCGCCCAAGTCTCAGCCAGACCTTAAGAGTCACTGAAACAGAAAGAG 2280
Db 2218 AAACATCGGAGGCGCCCAAGTCTCAGCCAGACCTTAAGAGTCACTGAAACAGAAAGAG 2277
QY 2281 CAGCACTATGAGAAAAGATTAAGTGTGGAACAATCAGATTAAGAAAAGACTGAGCTGAC 2340
Db 2278 CAGCACTATGAGAAAAGATTAAGTGTGGAACAATCAGATTAAGAAAAGACTGAGCTGAC 2337
QY 2341 AAGGAGACATGAGAGACATGATGCAAGACACAGAGAGAGGCCCATGAGAGGGCAAA 2400
Db 2338 AAGGAGACATGAGAGACATGATGCAAGACACAGAGAGAGGCCCATGAGAGGGCAAA 2397
QY 2401 AATCTCAGCGAAACGAAAGCGATGATCAATGCTATGGAATTCAGAAATGATCCCTGAA 2460
Db 2398 AATCTCAGCGAAACGAAAGCGATGATCAATGCTATGGAATTCAGAAATGATCCCTGAA 2457

QY 2461 CAGAGATTGTGAACTGTCTGAAAGCCAAATPAACTTGACAGAAATAGCACTCTTTTACC 2520
DB 2458 CAGAGATTGTGAACTGTCTGAAAGCCAAATPAACTTGACAGAAATAGCACTCTTTTACC 2517
QY 2521 CAAAGGAACATGAAAGGCCCAAGAAAGATGATTTCTGAACTTCAGGGCAACGAAATTTTAC 2580
DB 2518 CAAAGGAACATGAAAGGCCCAAGAAAGATGATTTCTGAACTTCAGGGCAACGAAATTTTAC 2577
QY 2581 CTGAGACACAGGCTGGGAAGTTGAGAGGCCAGAAACGAAAACTGAGAGAGCACTGGAG 2640
DB 2578 CTGAGACACAGGCTGGGAAGTTGAGAGGCCAGAAACGAAAACTGAGAGAGCACTGGAG 2637
QY 2641 AAGATCAGCCACCAAGACCAAGTGAACAAGATCGCTGTCTGAACTGAGACAAAGATTG 2700
DB 2638 AAGATCAGCCACCAAGACCAAGTGAACAAGATCGCTGTCTGAACTGAGACAAAGATTG 2697
QY 2701 CGGAGGTCAAGTCTGAGACAGAGGACCAAACTGAGAGCTCAAGGGCCAGCTCACAGAG 2760
DB 2698 CGGAGGTCAAGTCTGAGACAGAGGACCAAACTGAGAGCTCAAGGGCCAGCTCACAGAG 2757
QY 2761 CTACAGCTCTCCCTGACAGAGCGGAGTCAAGTTGACAGCCCTGACAGCTGACAGGGCG 2820
DB 2758 CTACAGCTCTCCCTGACAGAGCGGAGTCAAGTTGACAGCCCTGACAGCTGACAGGGCG 2817
QY 2821 GCCCTGAGAGCCAGCTTCCGACAGCGAAGACAGAGCTGGAAGAGACACAGCAAGAGCT 2880
DB 2818 GCCCTGAGAGCCAGCTTCCGACAGCGAAGACAGAGCTGGAAGAGACACAGCAAGAGCT 2877
QY 2881 GAAAGGAGATCCAGGCACTCAAGGCACTAGAGATGAAATCCAGGGCAATTTGAGTCT 2940
DB 2878 GAAAGGAGATCCAGGCACTCAAGGCACTAGAGATGAAATCCAGGGCAATTTGAGTCT 2937
QY 2941 CTTGCTAAGCTGTACTGTATATCAGAGCCTGGAGAGGAGCTPAAACAGCTGACCGAG 3000
DB 2938 CTTGCTAAGCTGTACTGTATATCAGAGCCTGGAGAGGAGCTPAAACAGCTGACCGAG 2997
QY 3001 GACACGCTGAATCTCAACCAAACTTCTTGTCCAAACAACTCGATGAGGCTTCT 3060
DB 2998 GACACGCTGAATCTCAACCAAACTTCTTGTGTCCAAACAACTCGATGAGGCTTCT 3057
QY 3061 GGGGCCAAGCAGATTTGTACAACTGCGAAATGAGTGAACCTTCCGCGGGAGATC 3120
DB 3058 GGGGCCAAGCAGATTTGTACAACTGCGAAATGAGTGAACCTTCCGCGGGAGATC 3117
QY 3121 ACCGAACGAGATGAGCTTTACAGAGCAGAGCAAGATGAGGCTGGAAGACCAAG 3180
DB 3118 ACCGAACGAGATGAGCTTTACAGAGCAGAGCAAGATGAGGCTGGAAGACCAAG 3177
QY 3181 TGCACCATGCTGAGAAACAGTTCATGATTTGAGAGCCCTPAAACGATGAGCTGTAGAA 3240
DB 3178 TGCACCATGCTGAGAAACAGTTCATGATTTGAGAGCCCTPAAACGATGAGCTGTAGAA 3237
QY 3241 AAAAGCGGCAAGTGGAGGCGCTGAGAGCGCTCTGGGTGATGAGAAATCCCACTTTGAG 3300
DB 3238 AAAAGCGGCAAGTGGAGGCGCTGAGAGCGCTCTGGGTGATGAGAAATCCCACTTTGAG 3297
QY 3301 TGTCCGTTGAGAGCTGAGAGGATGCTGGAACCGAGAAACGAGACAGAGCGAGAGCC 3360
DB 3298 TGTCCGTTGAGAGCTGAGAGGATGCTGGAACCGAGAAACGAGACAGAGCGAGAGCC 3357
QY 3361 GATCAGCGATCAACGAGTCTCGCCAGTGTGAGTGGCAGTGAAGAGACCAAGGCT 3420
DB 3358 GATCAGCGATCAACGAGTCTCGCCAGTGTGAGTGGCAGTGAAGAGACCAAGGCT 3417
QY 3421 GAGATTCTCGCTTGCAGCAGGCTCTCAAGAGCAAGAGCTGAAGCGCAGAGCTTCT 3480
DB 3418 GAGATTCTCGCTTGCAGCAGGCTCTCAAGAGCAAGAGCTGAAGCGCAGAGCTTCT 3477
QY 3481 GACAAGCTCAATGACCTGGAGAAAGAGATGCTATGCTGTAAGAAATGCCCCGAAGCTTA 3540
DB 3478 GACAAGCTCAATGACCTGGAGAAAGAGATGCTATGCTGTAAGAAATGCCCCGAAGCTTA 3537
QY 3541 CAGCAGAGCTGAGACTGAACGAGAGCTCAACAGAGGCTTCTGGAAGAGCAAGCCAA 3600

DB 3538 CAGCAGAGCTGAGACTGAACGAGAGCTCAACAGAGGCTTCTGGAAGAGCAAGCCAA 3597
QY 3601 TTACAGACAGATGAGACCTGACAGAAATCAATTTTCCGCTGACTCAAGACTGCA 3660
DB 3598 TTACAGACAGATGAGACCTGACAGAAATCAATTTTCCGCTGACTCAAGACTGCA 3657
QY 3661 GAAGCTCTAGATGGGCTGATCTACTGAAACAGAAAGAGTGAATTGAGATCACTG 3720
DB 3658 GAAGCTCTAGATGGGCTGATCTACTGAAACAGAAAGAGTGAATTGAGATCACTG 3717
QY 3721 GAAACATTTAGGTTCTTATTTCTCATGAAAGTGAATTTGAAAGGCACTATTTCTCA 3780
DB 3718 GAAACATTTAGGTTCTTATTTCTCATGAAAGTGAATTTGAAAGGCACTATTTCTCA 3777
QY 3781 CAACCAAACTCATTTGATTTCTGCAAGCCAAATGGAACCACTGCTAAAAAGAAAG 3840
DB 3778 CAACCAAACTCATTTGATTTCTGCAAGCCAAATGGAACCACTGCTAAAAAGAAAG 3837
QY 3841 GTTCTCTGAGTACAAATGAGCTGAAGCTGACCTTGGAGAAAGAAAGCTCGCTGCA 3900
DB 3838 GTTCTCTGAGTACAAATGAGCTGAAGCTGACCTTGGAGAAAGAAAGCTCGCTGCA 3897
QY 3901 GAGCTAGAGAAAGCCTTCAAGAAACCCGATGAGAGTCCGGTCCCGCGGAGGAGCT 3960
DB 3898 GAGCTAGAGAAAGCCTTCAAGAAACCCGATGAGAGTCCGGTCCCGCGGAGGAGCT 3957
QY 3961 GCCCAGCAGAAACAGAGCAAGCCACCAAGCTTCCAGCAGCAGCAGCAGGAGCAG 4020
DB 3958 GCCCAGCAGAAACAGAGCAAGCCACCAAGCTTCCAGCAGCAGCAGCAGGAGCAG 4017
QY 4021 ATGCGCATGTCCGCATCGTGCAGTGCAGAGACCAAGCCATGAGCTGAGCTG 4080
DB 4018 ATGCGCATGTCCGCATCGTGCAGTGCAGAGACCAAGCCATGAGCTGAGCTG 4077
QY 4081 GCCCGGCATCCAGCGCAGAAAGAGTCTTCACTCAAGAGAAATTAATGCGGCTT 4140
DB 4078 GCCCGGCATCCAGCGCAGAAAGAGTCTTCACTCAAGAGAAATTAATGCGGCTT 4137
QY 4141 AAGGAACGATGACCAACATATTCTCACCGATTCAACGTAGAGATGAAATGCGAGCC 4200
DB 4138 AAGGAACGATGACCAACATATTCTCACCGATTCAACGTAGAGATGAAATGCGAGCC 4197
QY 4201 ACAAGTGTCTGTGTCTGTGATACCGTGCATCTTGAACGCCAGCATCAATGTCTC 4260
DB 4198 ACAAGTGTCTGTGTCTGTGATACCGTGCATCTTGAACGCCAGCATCAATGTCTC 4257
QY 4261 GAATGTGAGTATGTGTCAACCCCAAGTCTCCACGTGCTGACGACCACTGCGGCTTG 4320
DB 4258 GAATGTGAGTATGTGTCAACCCCAAGTCTCCACGTGCTGACGACCACTGCGGCTTG 4317
QY 4321 CTTGCTGAATATGCCACACCTTCAACGAGGCTTCCGCGTGAACAAATGAACTCCCA 4380
DB 4318 CTTGCTGAATATGCCACACCTTCAACGAGGCTTCCGCGTGAACAAATGAACTCCCA 4377
QY 4381 GGTCTCCAGCAAGAGAGCCCAAGCAGCTTCACTGGAAGGAGTGAATGAGTCCC 4440
DB 4378 GGTCTCCAGCAAGAGAGCCCAAGCAGCTTCACTGGAAGGAGTGAATGAGTCCC 4437
QY 4441 AAGAAATPAAACAGAGACAGCAAGCTTGGACAGAGATCAATTTCTGAGAGATCA 4500
DB 4438 AAGAAATPAAACAGAGACAGCAAGCTTGGACAGAGATCAATTTCTGAGAGATCA 4497
QY 4501 AAGTCTCTATTTATGACATGAAAGCAGAGAGCTGGAACAGAGGCGGAGTGAATTT 4560
DB 4498 AAGTCTCTATTTATGACATGAAAGCAGAGAGCTGGAACAGAGGCGGAGTGAATTT 4557
QY 4561 GAGCTGTGCTTCCGAGCGGAGATGTATCTATTTCAATGAGTCCGTTGCTTCCGAATC 4620
DB 4558 GAGCTGTGCTTCCGAGCGGAGATGTATCTATTTCAATGAGTCCGTTGCTTCCGAATC 4617
QY 4621 GCAAAATPACAGCAAGCAGATGTCCATACATATCTGAAGATGAAATTTCAACCGCAGACC 4680

Db 4618 GCAAAATACAGCCAAAGCAGATGTCCATACATACAGATGAATCTCACCCGACACACC 4677
Qy 4681 ACCTGCTGGCCCGGAGAACTCTTACTTGTGTAAGTCTCCAGCTTCCCTGACAAACAGGCG 4740
Db 4678 ACCTGCTGGCCCGGAGAACTCTTACTTGTGTAAGTCTCCAGCTTCCCTGACAAACAGGCG 4737
Qy 4741 TGGGTCAACCGCTTAAATGATGATGTGCGAGGTGGAGAGTTTCTAGAGAAAGAGGAA 4800
Db 4738 TGGGTCAACCGCTTAAATGATGATGTGCGAGGTGGAGAGTTTCTAGAGAAAGAGGAA 4797
Qy 4801 GCTGATCTAAATGCTTGTGAAACTCCCTGTGAAACTGAAAGGTGATGACCGTCTAGAC 4860
Db 4798 GCTGATCTAAATGCTTGTGAAACTCCCTGTGAAACTGAAAGGTGATGACCGTCTAGAC 4857
Qy 4861 ATGAATCTGACGCTGCTTCTGATGACCAAGGTGTTGTGGCGACCGAGAGAGGCTC 4920
Db 4858 ATGAATCTGACGCTGCTTCTGATGACCAAGGTGTTGTGGCGACCGAGAGAGGCTC 4917
Qy 4921 TACGCCCTGATGTCTGAAAAACTCCCTAACCCATGTTCCAGAGAAATTGAGAGAGTCTTC 4980
Db 4918 TACGCCCTGATGTCTGAAAAACTCCCTAACCCATGTTCCAGAGAAATTGAGAGAGTCTTC 4977
Qy 4981 CAATTTATATATATCAAGACCTGAGAGAGTACTCATGATAGAGAGAGAGAGCGGCA 5040
Db 4978 CAATTTATATATATCAAGACCTGAGAGAGTACTCATGATAGAGAGAGAGAGCGGCA 5037
Qy 5041 CTGTGTCTGTGACGTGAGAAAGATGAAACAGTCCCTGCGCCAGTCCACCTGCGTCC 5100
Db 5038 CTGTGTCTGTGACGTGAGAAAGATGAAACAGTCCCTGCGCCAGTCCACCTGCGTCC 5097
Qy 5101 CACGCCGACATCTCACCCAACTTTTGAAGCTGTCAAGGGCTGGCCCTTGTGGGGCA 5160
Db 5098 CACGCCGACATCTCACCCAACTTTTGAAGCTGTCAAGGGCTGGCCCTTGTGGGGCA 5157
Qy 5161 GCGAAGATTGAGAACGGGCTCTGATCTGTGACGACCATGCGCAAAAGTCTGATCTTC 5220
Db 5158 GCGAAGATTGAGAACGGGCTCTGATCTGTGACGACCATGCGCAAAAGTCTGATCTTC 5217
Qy 5221 CGCTACACGAAACTCTGACAAATATCTGATCTCGGAAAGATGAGACCTGAGAGCC 5280
Db 5218 CGCTACACGAAACTCTGACAAATATCTGATCTCGGAAAGATGAGACCTGAGAGCC 5277
Qy 5281 TGCAGCTGATTCACCTTCCCAATTAAGTATCTCAATTTGGAACCAATTAATCTTACGA 5340
Db 5278 TGCAGCTGATTCACCTTCCCAATTAAGTATCTCAATTTGGAACCAATTAATCTTACGA 5337
Qy 5341 ATGACATGAGAGATGACGCTGAGAGATTCCTGATTAAGATGACATTCCTTGGCA 5400
Db 5338 ATGACATGAGAGATGACGCTGAGAGATTCCTGATTAAGATGACATTCCTTGGCA 5397
Qy 5401 CCGCTGTGTGTCGCGCTCTTCCAAACAGTTCCTGTCTCAATCGTGCAGGTGAACAG 5460
Db 5398 CCGCTGTGTGTCGCGCTCTTCCAAACAGTTCCTGTCTCAATCGTGCAGGTGAACAG 5457
Qy 5461 GCAAGGACAGAGAGATTAAGTGTGTTTCCAGAAATTGGAAGTTCGAGATTCT 6520
Db 5458 GCAAGGACAGAGAGATTAAGTGTGTTTCCAGAAATTGGAAGTTCGAGATTCT 5517
Qy 5521 TACGAAAGATGAGCGGACAGACGATCTCAAGTGAAGTCTTACTTGTGGCTTGGC 5580
Db 5518 TACGAAAGATGAGCGGACAGACGATCTCAAGTGAAGTCTTACTTGTGGCTTGGC 5577
Qy 5581 TACAGAGAAACCTATCTGTGTTGTGACCACTTCACTCACTGAAATTAATGATCCAG 5640
Db 5578 TACAGAGAAACCTATCTGTGTTGTGACCACTTCACTCACTGAAATTAATGATCCAG 5637
Qy 5641 GCAAGCTCTCAGAGAGAGCCCTGCGGAGGTACTTGAACATCCCGAACCGCGCTAC 5700
Db 5638 GCAAGCTCTCAGAGAGAGCCCTGCGGAGGTACTTGAACATCCCGAACCGCGCTAC 5697
Qy 5701 CTGGGCGCTGCAATTTCTCAGAGAGATTAAGTGGGTCCTCATACAGAGATTAATTA 5760
Db 5698 CTGGGCGCTGCAATTTCTCAGAGAGATTAAGTGGGTCCTCATACAGAGATTAATTA 5757

Qy 5761 AGGTCATTTGCTGCAAGGAAACTTGTGAAGAGTCCGCACTGAACCAACCGGCGC 5820
Db 5758 AGGTCATTTGCTGCAAGGAAACTTGTGAAGAGTCCGCACTGAACCAACCGGCGC 5817
Qy 5821 CCGTCCACCTCCCGACAGACCCCAACAGAGGCCCAACAGTCAACAGACATC 5880
Db 5818 CCGTCCACCTCCCGACAGACCCCAACAGAGGCCCAACAGTCAACAGACATC 5877
Qy 5881 ACCAAGGCGCTGAGCTTCAAGCCCAAGCGCCCGCAAGAGCCCAAGAGAGCCCA 5940
Db 5878 ACCAAGGCGCTGAGCTTCAAGCCCAAGCGCCCGCAAGAGCCCAAGAGAGCCCA 5937
Qy 5941 AGCAACCCCAAGCTTACCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6000
Db 5938 AGCAACCCCAAGCTTACCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5997
Qy 6001 CCGCCCTGAG 6060
Db 5998 CCGCCCTGAG 6057
Qy 6061 CCGCAGAGCTGTTGAAGACAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6120
Db 6058 CCGCAGAGCTGTTGAAGACAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6117
Qy 6121 CCGCTGTCCAGGTGAACAGGTCTGGAGACAGTCTTCAATATTA 6165
Db 6118 CCGCTGTCCAGGTGAACAGGTCTGGAGACAGTCTTCAATATTA 6162

RESULT 7
US-09-964-956-8
; Sequence 8, Application US/0964956
; Publication No. US20040043926A1
; GENERAL INFORMATION:
; APPLICANT: Gerlach, Valerie L
; APPLICANT: MacDougall, John R
; APPLICANT: Smithson, Glenda
; APPLICANT: Millet, Isabelle
; APPLICANT: Stone, David
; APPLICANT: Gunther, Erik
; APPLICANT: Ellerman, Karen
; APPLICANT: Grose, William M
; APPLICANT: Alsobrook II, John P
; APPLICANT: Lepley, Denise M
; APPLICANT: Burgess, Catherine B
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Spytek, Kimberly A
; APPLICANT: Leach, Martin D
; APPLICANT: Shimkets, Richard A
; TITLE OF INVENTION: No. US0040043926A1e1 Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-124
; CURRENT APPLICATION NUMBER: US/09/964, 956
; CURRENT FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: 60/235,631
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: 60/235,633
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: 60/235,808
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: 60/236,064
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: 60/236,065
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: 60/236,066
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: 60/236,135
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: 60/237,434
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/238,321
; PRIOR FILING DATE: 2000-10-05

```

; PRIOR APPLICATION NUMBER: 60/238,399
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/238,396
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/276,667
; PRIOR FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/294,823
; PRIOR FILING DATE: 2001-05-31
; PRIOR APPLICATION NUMBER: 60/304,868
; PRIOR FILING DATE: 2001-07-12
; NUMBER OF SEQ ID NOS: 127
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 8
; LENGTH: 6201
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-964-956-8

Query Match      99.0%; Score 6105.4; DB 11; Length 6201;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 6131; Conservative 0; Mismatches 21; Indels 3; Gaps 1;

QY      1  AATGTGAAGTTCAATATGAGAGCGGGAATCTTTGATGCTGCTGTAACCAT 60
DB      1  AATGTGAAGTTCAATATGAGAGCGGGAATCTTTGATGCTGCTGTAACCAT 60
QY      61  GCCAGCGGCGCTCCAGGCTGAATCTGTCTCCAGGGGAAACACCCCTTATGACTCA 120
DB      61  GCCAGCGGCGCTCCAGGCTGAATCTGTCTCCAGGGGAAACACCCCTTATGACTCA 120
QY      121  CACGAGATGTCCTCTTTCCGAGAGGATATAGATGCTCTTGTCTTTGAA 180
DB      121  CACGAGATGTCCTCTTTCCGAGAGGATATAGATGCTCTTGTCTTTGAA 180
QY      181  GAATGAGTGAAGCTGCTGATGAGATTAAGAGCTGAGCACTTTGCCGAAATAT 240
DB      181  GAATGAGTGAAGCTGCTGATGAGATTAAGAGCTGAGCACTTTGCCGAAATAT 240
QY      241  TCCGACACCATAGCTGATGATTAAGAGCTGAGCACTTTGCCGAAATAT 300
DB      241  TCCGACACCATAGCTGATGATTAAGAGCTGAGCACTTTGCCGAAATAT 300
QY      301  AGCTTGTAGCTGTGTGCTCACTTTGCTGAAGTGAAGTGTGAAGAGAAACCGGG 360
DB      301  AGCTTGTAGCTGTGTGCTCACTTTGCTGAAGTGAAGTGTGAAGAGAAACCGGG 360
QY      361  GAGATCTATGCTATGAGAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 420
DB      361  GAGATCTATGCTATGAGAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 420
QY      421  TTTTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 480
DB      421  TTTTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 480
QY      481  CAGTATGCTTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 540
DB      481  CAGTATGCTTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 540
QY      541  GACTTGTGCTCACTTTGAGATGATGAGAGAGAGAGAGAGAGAGAGAGAGAT 600
DB      541  GACTTGTGCTCACTTTGAGATGATGAGAGAGAGAGAGAGAGAGAGAGAGAT 600
QY      601  TACTTGTGCTGAGCTATTTGGCTGCTCAAGAGCTTATGATGAGATGAGATG 660
DB      601  TACTTGTGCTGAGCTATTTGGCTGCTCAAGAGCTTATGATGAGATGAGATG 660
QY      661  GAGATCAAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 720
DB      661  GAGATCAAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 720
QY      721  GAGATCTGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 780
DB      721  GAGATCTGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 780
QY      777  GAGATCTGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 777
DB      777  GAGATCTGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 777
```

```

QY      781  CGAGATTACATGCTCTCTGAGAGTCTGATGATGAGAGAGAGAGAGAGAG 840
DB      778  CGAGATTACATGCTCTCTGAGAGTCTGATGATGAGAGAGAGAGAGAGAG 837
QY      841  GGCCTGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 900
DB      838  GGCCTGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 897
QY      901  TCCCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 960
DB      898  TCCCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 957
QY      961  TTTTGAATTTTCAAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1020
DB      958  TTTTGAATTTTCAAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1017
QY      1021  TTGTTGTGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1080
DB      1018  TTGTTGTGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1077
QY      1081  TCTAAATTTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1140
DB      1078  TCTAAATTTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1137
QY      1141  TCTGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1200
DB      1138  TCTGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1197
QY      1201  CCGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1260
DB      1198  CCGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1257
QY      1261  TACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1320
DB      1258  TACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1317
QY      1321  CTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1380
DB      1318  CTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1377
QY      1381  TCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1440
DB      1378  TCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1437
QY      1441  GAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1500
DB      1438  GAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1497
QY      1501  TCCCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1560
DB      1498  TCCCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1557
QY      1561  TTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1620
DB      1558  TTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1617
QY      1621  GATTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1680
DB      1618  GATTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1677
QY      1681  GTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1740
DB      1678  GTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1737
QY      1741  CGAGAGATTTCTAAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1800
DB      1738  CGAGAGATTTCTAAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1797
QY      1801  CGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1860
DB      1798  CGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1857
```

QY 1861 GTGGAGAAATATGGGAACTGGAAGATCAATGCTGAGCAGACGCTCAAAATTCAGAG 1920
| | | | |
Db 1858 GTGGAGAAATATGGGAACTGGAAGATCAATGCTGAGCAGACGCTCAAAATTCAGAG 1917
| | | | |
QY 1921 CTCGAAGAGAACTGGAAGAGCTGTAAACCCAGCACTGAGGCACTGAGCTGCTGAG 1980
| | | | |
Db 1918 CTCGAAGAGAACTGGAAGAGCTGTAAACCCAGCACTGAGGCACTGAGCTGCTGAG 1977
| | | | |
QY 1981 AATTTCGCCGAGGCAAGAGGAGGCGAGAGGAGGCTGAGAGAGCTGAGAGCCGAGAG 2040
| | | | |
Db 1978 AATTTCGCCGAGGCAAGAGGAGGCGAGAGGAGGCTGAGAGAGCTGAGAGCCGAGAG 2037
| | | | |
QY 2041 GATTCTTCTGAAGGATCAGAAAGAGCTGAGAGGCTGAGAGAGCCGCGCATTCCTG 2100
| | | | |
Db 2038 GATTCTTCTGAAGGATCAGAAAGAGCTGAGAGGCTGAGAGCCGCGCATTCCTG 2097
| | | | |
QY 2101 GAGAACAGGTAAAGAGCTAGAGACCATGAGCGTAGAGAAAACAGACTGAAGATGAC 2160
| | | | |
Db 2098 GAGAACAGGTAAAGAGCTAGAGACCATGAGCGTAGAGAAAACAGACTGAAGATGAC 2157
| | | | |
QY 2161 ATCCAGCAAAATCCCAACAGATCAGACAGATGCTGATAAATTTCTGAGCTGAAAG 2220
| | | | |
Db 2158 ATCCAGCAAAATCCCAACAGATCAGACAGATGCTGATAAATTTCTGAGCTGAAAG 2217
| | | | |
QY 2221 AAACATGAGGAGGCCAAAGTCTCAGCCAGACCTAGAGTGCCTGAAACAGAAAGAG 2280
| | | | |
Db 2218 AAACATGAGGAGGCCAAAGTCTCAGCCAGACCTAGAGTGCCTGAAACAGAAAGAG 2277
| | | | |
QY 2281 CAGCACTATGAGAAAAGATTAAAGTTTGAACATCAGATGAAGAAAGACTGAGTAC 2340
| | | | |
Db 2278 CAGCACTATGAGAAAAGATTAAAGTTTGAACATCAGATGAAGAAAGACTGAGTAC 2337
| | | | |
QY 2341 AAGGAGACATGAGAACATGATGACAGACACAGAGAGAGGCCATGAGAGGCGAAA 2400
| | | | |
Db 2338 AAGGAGACATGAGAACATGATGACAGACACAGAGAGAGGCCATGAGAGGCGAAA 2397
| | | | |
QY 2401 ATTCTGAGGAAACAGAGGCGATGATCAATGCTATGATTCGAAGATCAGATCCCTGAA 2460
| | | | |
Db 2398 ATTCTGAGGAAACAGAGGCGATGATCAATGCTATGATTCGAAGATCAGATCCCTGAA 2457
| | | | |
QY 2461 CAGAGATTTGTGAACTGTCTGAAGCCAAATACTTGACAGCAATAGCAGTCTTTTAC 2520
| | | | |
Db 2458 CAGAGATTTGTGAACTGTCTGAAGCCAAATACTTGACAGCAATAGCAGTCTTTTAC 2517
| | | | |
QY 2521 CAAAGAACATGAGAGGCCCAAGAAAGATGATTTCTGAACCTCAGGCAACGAAATTTTAC 2580
| | | | |
Db 2518 CAAAGAACATGAGAGGCCCAAGAAAGATGATTTCTGAACCTCAGGCAACGAAATTTTAC 2577
| | | | |
QY 2581 CTGGAGACACAGGCTGGGAAGTTGGAGGCCCAAGAACCTGAGAGAGCTGGAG 2640
| | | | |
Db 2578 CTGGAGACACAGGCTGGGAAGTTGGAGGCCCAAGAACCTGAGAGAGCTGGAG 2637
| | | | |
QY 2641 AAGATCAGCACAACAGACCAACAGTGAACAAGATCGGCTGCTGAGACTGAGAGCAAGTTG 2700
| | | | |
Db 2638 AAGATCAGCACAACAGACCAACAGTGAACAAGATCGGCTGCTGAGACTGAGAGCAAGTTG 2697
| | | | |
QY 2701 CGGAGGTCAGTCTAGACACAGAGCAGAAACTGAGACTCAAGGCCAGCTCACAAG 2760
| | | | |
Db 2698 CGGAGGTCAGTCTAGACACAGAGCAGAAACTGAGACTCAAGGCCAGCTCACAAG 2757
| | | | |
QY 2761 CTACAGCTCTCCCTGAGAGAGGCGGAGTGCAGTTGACAGCCCTGAGGCTCACAAG 2820
| | | | |
Db 2758 CTACAGCTCTCCCTGAGAGAGGCGGAGTGCAGTTGACAGCCCTGAGGCTCACAAG 2817
| | | | |
QY 2821 GCCCTGAGAGGCGAGCTTGCCAGGAGAGACAGAGCTGAGAGACCAAGAGAGCT 2880
| | | | |
Db 2818 GCCCTGAGAGGCGAGCTTGCCAGGAGAGACAGAGCTGAGAGACCAAGAGAGCT 2877
| | | | |
QY 2881 GAAAGAGAGATCAGGCACTCAACGCAATAGAGTGAATCAGGCAATTTGATGCT 2940
| | | | |
Db 2878 GAAAGAGAGATCAGGCACTCAACGCAATAGAGTGAATCAGGCAATTTGATGCT 2937
| | | | |
QY 2941 CTTGCTAACAGCTGATCTATATCAAGACTGGAGAGAGCTTAAACAGCTGACCGAG 3000
| | | | |

Db 2938 CTTGCTAACAGCTGATCTGATCAAGACCTGAGAGAGCTTAAACAGCTGACCGAG 2997
| | | | |
QY 3001 GACAACTGAACTCAACAAACCAAACTTCTGATGCTCAAACTGATGAGCTTCT 3060
| | | | |
Db 2998 GACAACTGAACTCAACAAACCAAACTTCTGATGCTCAAACTGATGAGCTTCT 3057
| | | | |
QY 3061 GCGCCACGACGAGATTGTACAACTGCGAAGTGAAGTGAACCATCTCCGCGGAGATC 3120
| | | | |
Db 3058 GCGCCACGACGAGATTGTACAACTGCGAAGTGAAGTGAACCATCTCCGCGGAGATC 3117
| | | | |
QY 3121 ACGGAACGAGATGACGCTTACACGACCAAGCAACATGAGAGCTTGAAGACACG 3180
| | | | |
Db 3118 ACGGAACGAGATGACGCTTACACGACCAAGCAACATGAGAGCTTGAAGACACG 3177
| | | | |
QY 3181 TGCAACATGCTGAGAGAACAGCTCATGGAATTTGAGGCGCTTAAACGATGAGCTGAGAA 3240
| | | | |
Db 3178 TGCAACATGCTGAGAGAACAGCTCATGGAATTTGAGGCGCTTAAACGATGAGCTGAGAA 3237
| | | | |
QY 3241 AAAGAGCGGAGTGGAGGCGCTGAGAGAGCGCTCTGGGTGATGAGAAATCCAGTTTGA 3300
| | | | |
Db 3238 AAAGAGCGGAGTGGAGGCGCTGAGAGAGCGCTCTGGGTGATGAGAAATCCAGTTTGA 3297
| | | | |
QY 3301 TGTCCGCTTGGAGAGCTGACAGAGATGCTGAGACCCGAGAAACAGAGCGGCGAGAGCC 3360
| | | | |
Db 3298 TGTCCGCTTGGAGAGCTGACAGAGATGCTGAGACCCGAGAAACAGAGCGGCGAGAGCC 3357
| | | | |
QY 3361 GATCAGCGGATCAACCGAGTCTGCGCAGAGTGGTGAAGCTGAGTGAAGAGCAAGGCT 3420
| | | | |
Db 3358 GATCAGCGGATCAACCGAGTCTGCGCAGAGTGGTGAAGCTGAGTGAAGAGCAAGGCT 3417
| | | | |
QY 3421 GAGATTTCTGCTTGGAGAGGCTCTCAAGAGCAGAGCTGAAGGCCGAGAGCTCTCT 3480
| | | | |
Db 3418 GAGATTTCTGCTTGGAGAGGCTCTCAAGAGCAGAGCTGAAGGCCGAGAGCTCTCT 3477
| | | | |
QY 3481 GACAACTGAACTGACCTGAGAGAAAGCAATGCTATGCTTGAATGAATGCCGAACTTA 3540
| | | | |
Db 3478 GACAACTGAACTGACCTGAGAGAAAGCAATGCTATGCTTGAATGAATGCCGAACTTA 3537
| | | | |
QY 3541 CAGCAGAACTGAGAGCTGAAGAGAGCTCAACAGAGGCTCTGAGAGAGCAAGCCAA 3600
| | | | |
Db 3538 CAGCAGAACTGAGAGCTGAAGAGAGCTCAACAGAGGCTCTGAGAGAGCAAGCCAA 3597
| | | | |
QY 3601 TTACAGCAGCAGATGAGCTGAGAAAATCAATTTTCGCTGACTCAAGGACTGCA 3660
| | | | |
Db 3598 TTACAGCAGCAGATGAGCTGAGAAAATCAATTTTCGCTGACTCAAGGACTGCA 3657
| | | | |
QY 3661 GAACTCTAGATCGGCTGATCTAAGAACAGAAAGATGACTTGGAGATTCAGCTG 3720
| | | | |
Db 3658 GAACTCTAGATCGGCTGATCTAAGAACAGAAAGATGACTTGGAGATTCAGCTG 3717
| | | | |
QY 3721 GAAACATTCAGGTTCTGATCTCATGAAAGGTGAAGTGAAGGCGACTATTTCTGAA 3780
| | | | |
Db 3718 GAAACATTCAGGTTCTGATCTCATGAAAGGTGAAGTGAAGGCGACTATTTCTGAA 3777
| | | | |
QY 3781 CAAACCAAACTGATGATTTTCTGCAAGCCAAATGGAACCACTGCTTAAAAAGAAAAG 3840
| | | | |
Db 3778 CAAACCAAACTGATGATTTTCTGCAAGCCAAATGGAACCACTGCTTAAAAAGAAAAG 3837
| | | | |
QY 3841 GTTCTCTGCACTAAGAGCTGAAAGCTGGCCCTGAGAGAGAGAAAGCTGCTGTGCA 3900
| | | | |
Db 3838 GTTCTCTGCACTAAGAGCTGAAAGCTGGCCCTGAGAGAGAGAAAGCTGCTGTGCA 3897
| | | | |
QY 3901 GAGCTAGAGAGAGCCCTCAGAGAGCCGCAATCGAGCTCCGGTCCGCCCGGAGGAGACT 3960
| | | | |
Db 3898 GAGCTAGAGAGAGCCCTCAGAGAGCCGCAATCGAGCTCCGGTCCGCCCGGAGGAGACT 3957
| | | | |
QY 3961 GCCCACCGCAAGCAACGAGCAACCCACATCCATCCAGCCAGCCAGCGAGGAGAG 4020
| | | | |
Db 3958 GCCCACCGCAAGCAACGAGCAACCCACATCCATCCAGCCAGCGAGGAGAG 4017
| | | | |
QY 4021 ATGCACTGTCGCGCATGCTGTGCGGCTGCGCAGAGACCAAGCCAGTGCATGAGCTGCTG 4080
| | | | |

Dh 4018 ATGCCATGTCGTGCGATCGTGGGTGCGCAGAGACCAAGCCAGTGCATGAGCCTGCTG 4077
Qy 4081 GCCCGGCATCCAGCGCGAGAAAGAGTCTTCAATCCAGAGGAAATTAGTCGGCTCTT 4140
Dh 4078 GCCCGGCATCCAGCGCGAGAAAGAGTCTTCAATCCAGAGGAAATTAGTCGGCTCTT 4137
Qy 4141 AAGGAAGCATGACGACCAATATTCTCTACCGATTCAACGTAGACTGAACATGCGAGCC 4200
Dh 4138 AAGGAAGCATGACGACCAATATTCTCTACCGATTCAACGTAGACTGAACATGCGAGCC 4197
Qy 4201 ACAAAGTGTCTGTGTCTGTGATACCTGTGACCTTTGAGCGCGGATCCAAATGTCTC 4260
Dh 4198 ACAAAGTGTCTGTGTCTGTGATACCTGTGACCTTTGAGCGCGGATCCAAATGTCTC 4257
Qy 4261 GAATGTGAGGTGATGTGTGACCCCAAGTGTCTGCGCTTGCAGGCACTGTGGGCTTG 4320
Dh 4258 GAATGTGAGGTGATGTGTGACCCCAAGTGTCTGCGCTTGCAGGCACTGTGGGCTTG 4317
Qy 4321 CTGTGTAATATGCGACACATTCACCGAGGCTTGTGCGGTGCAAAATGAACTCCCA 4380
Dh 4318 CTGTGTAATATGCGACACATTCACCGAGGCTTGTGCGGTGCAAAATGAACTCCCA 4377
Qy 4381 GGTCTCCAGACCAAGAGGCCAGAGACGACTTGGACCTTGAAGGTGTGAAGTGTCCC 4440
Dh 4378 GGTCTCCAGACCAAGAGGCCAGAGACGACTTGGACCTTGAAGGTGTGAAGTGTCCC 4437
Qy 4441 AGGAATTAACAAACGAGGACAGCAAGCTGGGACAGAAATGATCAATTGCTCTGAGGATCA 4500
Dh 4438 AGGAATTAACAAACGAGGACAGCAAGCTGGGACAGAAATGATCAATTGCTCTGAGGATCA 4497
Qy 4501 AAGTCTCTATTTATGACAAATGAAGCCAGAGAACTTGACAGAGGCCGTGGAAGAAATT 4560
Dh 4498 AAGTCTCTATTTATGACAAATGAAGCCAGAGAACTTGACAGAGGCCGTGGAAGAAATT 4557
Qy 4561 GAGTGTGCTTCTCCGACGGGGATGTATCTATTATGTGTGCTGTGGTCTCCGAATCTC 4620
Dh 4558 GAGTGTGCTTCTCCGACGGGGATGTATCTATTATGTGTGCTGTGGTCTCCGAATCTC 4617
Qy 4621 GCAAAATCAGCCAAAGAGATGTCCCATACATCTGAAGATGAACTCACCCGACACCC 4680
Dh 4618 GCAAAATCAGCCAAAGAGATGTCCCATACATCTGAAGATGAACTCACCCGACACCC 4677
Qy 4681 ACCGTGTGCGCCGGGAGAACCTCTACTGTAGCTCCAGCTTCCCTGACAAACAGGCG 4740
Dh 4678 ACCGTGTGCGCCGGGAGAACCTCTACTGTAGCTCCAGCTTCCCTGACAAACAGGCG 4737
Qy 4741 TGGGTCAACCGCTTGAATCAATGTTGTGCGAGGTGGAGAGTTTCTAGGAAAAAGCAGAA 4800
Dh 4738 TGGGTCAACCGCTTGAATCAATGTTGTGCGAGGTGGAGAGTTTCTAGGAAAAAGCAGAA 4797
Qy 4801 GCTGATGCTAAACGTGTGGAACCTCCCTGTGAAACCTGGAAGGTGATGACCGTCTAGAC 4860
Dh 4798 GCTGATGCTAAACGTGTGGAACCTCCCTGTGAAACCTGGAAGGTGATGACCGTCTAGAC 4857
Qy 4861 ATGAATGACGCTGCTTCACTGACGAGGTGTGTGTGGCACCAGGAGAGGGCTC 4920
Dh 4858 ATGAATGACGCTGCTTCACTGACGAGGTGTGTGTGGCACCAGGAGAGGGCTC 4917
Qy 4921 TAGCGCTGAATGTCTTGAAGAACTCCCTTAACCAATGCTCCAGAAATGGACAGTCTTC 4980
Dh 4918 TAGCGCTGAATGTCTTGAAGAACTCCCTTAACCAATGCTCCAGAAATGGACAGTCTTC 4977
Qy 4981 CAAATTTATATTATCAAGGACCTGGAAGAGTACATGATAGCAGAGAGAGAGGGCA 5040
Dh 4978 CAAATTTATATTATCAAGGACCTGGAAGAGTACATGATAGCAGAGAGAGAGGGCA 5037
Qy 5041 CTGTGTCTTGTGACGTTGAAGAAAGTGAACAGTCCCTGCGCCAGTCCCACTGCTGCTCC 5100
Dh 5038 CTGTGTCTTGTGACGTTGAAGAAAGTGAACAGTCCCTGCGCCAGTCCCACTGCTGCTCC 5097
Qy 5101 CAGCCGACATCTACCCCAACATTTTGAAGCTGTCAAGGGCTGCACTTGTGTGGGCA 5160
Dh 5098 CAGCCGACATCTACCCCAACATTTTGAAGCTGTCAAGGGCTGCACTTGTGTGGGCA 5157

Qy 5161 GGCAGATTGAGAACGGGCTCTGATCTGTGACGCCATGCGCCAGAAAGTGTCTATCTC 5220
Dh 5158 GGCAGATTGAGAACGGGCTCTGATCTGTGACGCCATGCGCCAGAAAGTGTCTATCTC 5217
Qy 5221 CGCTACACGAAAAACCTCAGCAAAATATGCAATCCGAAAGAGATGAGACTCAGAGCC 5280
Dh 5218 CGCTACACGAAAAACCTCAGCAAAATATGCAATCCGAAAGAGATGAGACTCAGAGCC 5277
Qy 5281 TGCAGCTGTATCCACTTCACTCAATTAAGATCTCTCAATTGGAACCAATTAATCTAGAA 5340
Dh 5278 TGCAGCTGTATCCACTTCACTCAATTAAGATCTCTCAATTGGAACCAATTAATCTAGAA 5337
Qy 5341 ATGACATGAAGAGATCAACGCTCGAAGAAATCTGTGATTAAGATCAATCTCTGCGCA 5400
Dh 5338 ATGACATGAAGAGATCAACGCTCGAAGAAATCTGTGATTAAGATCAATCTCTGCGCA 5397
Qy 5401 CTGTGTGTGTTTCCCGCTTCTTCAACAGTTCCTGTCTCATGCGGAGGTGAAGAGC 5460
Dh 5398 CTGTGTGTGTTTCCCGCTTCTTCAACAGTTCCTGTCTCATGCGGAGGTGAAGAGC 5457
Qy 5461 GCAGGGCAGGAGAGAGTACTTGTGTGTTCACGAAATTTGAGAGTGTGTTGTTCT 5520
Dh 5458 GCAGGGCAGGAGAGAGTACTTGTGTGTTCACGAAATTTGAGAGTGTGTTGTTCT 5517
Qy 5521 TAGGAAAGAGTACGCGCACAGAGATCTCAAGTGAAGTGTGCTTACCTTTGGCTTTGCC 5580
Dh 5518 TAGGAAAGAGTACGCGCACAGAGATCTCAAGTGAAGTGTGCTTACCTTTGGCTTTGCC 5577
Qy 5581 TACAGAGAACCTCATCTGTTTGTGACCCACTTCACTCACTCCAGAAATTTGAGATCCAG 5640
Dh 5578 TACAGAGAACCTCATCTGTTTGTGACCCACTTCACTCACTCCAGAAATTTGAGATCCAG 5637
Qy 5641 GCAAGCTCTCAGAGAGGACCCCTGCGCGAGAGTACTGTGACATCCCGAGCCCGCTAC 5700
Dh 5638 GCAAGCTCTCAGAGAGGACCCCTGCGCGAGAGTACTGTGACATCCCGAGCCCGCTAC 5697
Qy 5701 CTGGGCGCTGCCATTTCTCAGAGAGGATTTACTTGGCTGTCTCATACAGATTAATTA 5760
Dh 5698 CTGGGCGCTGCCATTTCTCAGAGAGGATTTACTTGGCTGTCTCATACAGATTAATTA 5757
Qy 5761 AGGTCATTTGTGTGCAAGGAGAACTGTGTAAGAGTGTGCGCACTGAACACCAAGGGGCG 5820
Dh 5758 AGGTCATTTGTGTGCAAGGAGAACTGTGTAAGAGTGTGCGCACTGAACACCAAGGGGCG 5817
Qy 5821 CCGTCCACCTCCGACAGCCCAACAGAGAGGCCCAACCAAGTCAACAGACATC 5880
Dh 5818 CCGTCCACCTCCGACAGCCCAACAGAGAGGCCCAACCAAGTCAACAGACATC 5877
Qy 5881 ACCAAGCGGTGCGCTCAGCCAGCGCGCCGAGAGGCCCAAGCCCAAGCCCAAGGAGCA 5940
Dh 5878 ACCAAGCGGTGCGCTCAGCCAGCGCGCCGAGAGGCCCAAGCCCAAGGAGCA 5937
Qy 5941 AGCACACCCCAACCGCTACCGGAGGGGCGAGACGAGTGTGCGAGAGGCAAGTCTCTGGC 6000
Dh 5938 AGCACACCCCAACCGCTACCGGAGGGGCGAGACGAGTGTGCGAGAGGCAAGTCTCTGGC 5997
Qy 6001 CGCCCGCTGAGAGGAGAGTCCCGCGCGAGTACTCAGACAGCGGAGAGAGCGGTCC 6060
Dh 5998 CGCCCGCTGAGAGGAGAGTCCCGCGCGAGTACTCAGACAGCGGAGAGAGCGGTCC 6057
Qy 6061 CCCGCGAGGCTGTTTGAAGACAGAGAGGGGCCGCTGCGTGGGAGCCGTGAGAGCC 6120
Dh 6058 CCCGCGAGGCTGTTTGAAGACAGAGAGGGGCCGCTGCGTGGGAGCCGTGAGAGCC 6117
Qy 6121 CCGCTGTCCAGGTGAACAGAGTCTGGAGACAGTCC 6155
Dh 6118 CCGCTGTCCAGGTGAACAGAGTCTGGAGACAGTCC 6152

RESULT 8
US-10-262-511-13
; Sequence 13, Application US/10262511

Query Match	99.0%	Score 6105.4	DB 17	Length 6201
Best Local Similarity	99.6%	Pred. No. 0		
Matches 6131	Conservative 0	Mismatches 21	Indels 3	Gaps 1

QY	1	ATGTGAGTTCATATATGAGCGCGGAATCTTTGATGCTGCGTCGCTGAACCAATT	60

D	b	1	ATGTGAAGTTCAAAATATGAGAGCGGAAATCCTTTGATCTGTGCTGCTGAACCAATT	60
Q	y	61	GCCAGCGGGCTTCCAGGCTGAATCTGTTCCTCCAGGGGAAACCACTTTATGACTCAA	120
D	b	61	GCCAGCGGGCTTCCAGGCTGAATCTGTTCCTCCAGGGGAAACCACTTTATGACTCAA	120
Q	y	121	CAGAGATGTCTCTCTTTTCCGAGAAAGGAAATTAGATCCCTCTTGTTCCTTTGAA	180
D	b	121	CACAGATGTCTCTCTTTTCCGAGAAAGGAAATTAGATCCCTCTTGTTCCTTTGAA	180
Q	y	181	GAATGCAGTAGAGCTGCTGTATGAAGAATTAAACAAGTAGCAACTTGTCCGGAAAGTAT	240
D	b	181	GAATGCAGTAGAGCTGCTGTATGAAGAATTAAACAAGTAGCAACTTGTCCGGAAAGTAT	240
Q	y	241	TCCGACACCAATAGCTGATTAACAGAGCTCCAGCTTCGGCAAAAGACTTCGAAGTCAGA	300
D	b	241	TCCGACACCAATAGCTGATTAACAGAGCTTCAGAGCTTCGGCAAAAGACTTCGAAGTCAGA	300
Q	y	301	AGCTTGTAGTGTGTGTCTCACTTTGCTGAAGTGCAGGTGGTAAAGAGAAAGCAACCGGG	360
D	b	301	AGCTTGTAGTGTGTGTGTCTCACTTTGCTGAAGTGCAGGTGGTAAAGAGAAAGCAACCGGG	360
Q	y	361	GACATCTATGCTATGAAGATGTAAGAAAGAAAGGCTTTTATGGCCCAAGACAGGTTTCA	420
D	b	361	GACATCTATGCTATGAAGATGTAAGAAAGAAAGGCTTTTATGGCCCAAGACAGGTTTCA	420
Q	y	421	TTTTTTGAGAAAGCGGAAACATATTATCTCGAAGACAAAGCCCGTGATATCCCCCAATTA	480
D	b	421	TTTTTTGAGAAAGCGGAAACATATTATCTCGAAGACAAAGCCCGTGATATCCCCCAATTA	480
Q	y	481	CAGTATGCTTTCCAGACAAAATCACTTTATCTGTCAATGAATATCAGCTGTGAGGG	540
D	b	481	CAGTATGCTTTCCAGACAAAATCACTTTATCTGTCAATGAATATCAGCTGTGAGGG	540
Q	y	541	GACTTGCTGTCACTTTGAAATGATATGATAGGACCAAGTTAGAGAAACCTGATACAGTTT	600
D	b	541	GACTTGCTGTCACTTTGAAATGATATGATAGGACCAAGTTAGAGAAACCTGATACAGTTT	600
Q	y	601	TACCTAGCTGAGCTGATTTTGGCTGTTCACAGCGTTCACTGATAGGATATACGTGATCGA	660
D	b	601	TACCTAGCTGAGCTGATTTTGGCTGTTCACAGCGTTCACTGATAGGATATACGTGATCGA	660
Q	y	661	GACATCAAGCTGAGAACATTTCTGTTGACCGCACAGGACACATCAAGCTGTGTGATTTT	720
D	b	661	GACATCAAGCTGAGAACATTTCTGTTGACCGCACAGGACACATCAAGCTGTGTGATTTT	720
Q	y	721	GGATGTGCGGGGAAATGATTTCAAACAAGATGTGATATGCCAACTCCGATTTGGAGCC	780
D	b	721	GGATGTGCGGGGAAATGATTTCAAACAAGATGTGATATGCCAACTCCGATTTGGAGCC	780
Q	y	781	CCAGATTACATGAGCTCTGTAAGTGTGACTGTGATGAACGGGGATGGAAGGACCTTAC	840
D	b	781	CCAGATTACATGAGCTCTGTAAGTGTGACTGTGATGAACGGGGATGGAAGGACCTTAC	840
Q	y	841	GACCTGAGCTGTGACTGTGTGCTGAGTGTGGCGCTGATTTGCTATGAGATTTTATGGAGGA	900
D	b	841	GACCTGAGCTGTGACTGTGTGCTGAGTGTGGCGCTGATTTGCTATGAGATTTTATGGAGGA	900
Q	y	901	TCCCCCTTCCGAGAGGGAACCTCTGCGCAAACTTCAATAACATTATGAAATTTCCAGCGG	960
D	b	901	TCCCCCTTCCGAGAGGGAACCTCTGCGCAAACTTCAATAACATTATGAAATTTCCAGCGG	960
Q	y	961	TTTTTGAATTTCCAGATGACCCCAAAAGTAGACGTACTTTCTTGTATCTGATTTCAAGC	1020
D	b	961	TTTTTGAATTTCCAGATGACCCCAAAAGTAGACGTACTTTCTTGTATCTGATTTCAAGC	1020
Q	y	1021	TTGTGTGCGGCCGAAAGAGAGACTGAAGTTTGAAGGTCTTTGCTCCCAATCCTTTCTTC	1080
D	b	1021	TTGTGTGCGGCCGAAAGAGAGACTGAAGTTTGAAGGTCTTTGCTCCCAATCCTTTCTTC	1080
Q	y	1081	TTCTAAATTTGACTGGAACCAATTCTGTAATCTCTCTCCCTCTTGTCTTCCACCTCTAAG	1140
D	b	1081	TTCTAAATTTGACTGGAACCAATTCTGTAATCTCTCTCCCTCTTGTCTTCCACCTCTAAG	1140
Q	y	1140	TTCTAAATTTGACTGGAACCAATTCTGTAATCTCTCTCCCTCTTGTCTTCCACCTCTAAG	1140
D	b	1140	TTCTAAATTTGACTGGAACCAATTCTGTAATCTCTCTCCCTCTTGTCTTCCACCTCTAAG	1140

QY 1141 TCTGACGATGACACCTTCCAAATTTTGTATGAACAGAGAAATTCGTGGTTTCATCTCT 1200
DB 1138 TCTGACGATGACACCTTCCAAATTTGTATGAACAGAGAAATTCGTGGTTTCATCTCT 1197
QY 1201 CCGGTCCAGCTGAGCCCCCTCAGGCTTCTCGGGTGAAGAACTGCCGTTTGTGGGGTTTTCG 1260
DB 1198 CCGGTCCAGCTGAGCCCCCTCAGGCTTCTCGGGTGAAGAACTGCCGTTTGTGGGGTTTTCG 1257
QY 1261 TACAGCAAGGCACTGGGGATTCCTTGTGTAGATCTGAGTCTGTGTGTGTGCGGCTTGAATCC 1320
DB 1258 TACAGCAAGGCACTGGGGATTCCTTGTGTAGATCTGAGTCTGTGTGTGTGCGGCTTGAATCC 1317
QY 1321 CTTGCCAAGACTAGCTCCATGGAAGAAAGAACTTCTCATCAAAAGCAAAAGACTCAAGAC 1380
DB 1318 CTTGCCAAGACTAGCTCCATGGAAGAAAGAACTTCTCATCAAAAGCAAAAGACTCAAGAC 1377
QY 1381 TCTCAGGACAAAGTGTCAACAGATGAGACGAGAAATGACCCCGTTACATCCGAGAGTGTCA 1440
DB 1378 TCTCAGGACAAAGTGTCAACAGATGAGACGAGAAATGACCCCGTTACATCCGAGAGTGTCA 1437
QY 1441 GAGGTGAGGCTGTAGTCTTGTAGTCAAGAGAGGTGAGGCTGAGGCTCTGAGACTCAGAGA 1500
DB 1438 GAGGTGAGGCTGTAGTCTTGTAGTCAAGAGAGGTGAGGCTGAGGCTCTGAGACTCAGAGA 1497
QY 1501 TCCCTCCTGAGCAGAGACTTGTCTACCTACATCAACAGATGCAAGTACTTAAAGCGAAGT 1560
DB 1498 TCCCTCCTGAGCAGAGACTTGTCTACCTACATCAACAGATGCAAGTACTTAAAGCGAAGT 1557
QY 1561 TTGAGCAAGACAGATGAGAGTGTCCAGAGAGATGACAAAGCACTGACGCTTCTCAT 1620
DB 1558 TTGAGCAAGACAGATGAGAGTGTCCAGAGAGATGACAAAGCACTGACGCTTCTCAT 1617
QY 1621 GATTTCAGAGCGAGCCGGAAGCTCCAGAAATTCAAAGAGCGAGTACCAAGCTCAA 1680
DB 1618 GATTTCAGAGCGAGCCGGAAGCTCCAGAAATTCAAAGAGCGAGTACCAAGCTCAA 1677
QY 1681 GTGGAAGAAATGAGTGTGATGATGAATCAAGTGGAGAGATCTTGTCTCAGCAAGAGA 1740
DB 1678 GTGGAAGAAATGAGTGTGATGATGAATCAAGTGGAGAGATCTTGTCTCAGCAAGAGA 1737
QY 1741 CGAGTGAATCTTACGAATCTGAGCTGAGAGATGTCCGGCTTGTCTGTAAGAAATTCAG 1800
DB 1738 CGAGTGAATCTTACGAATCTGAGCTGAGAGATGTCCGGCTTGTCTGTAAGAAATTCAG 1797
QY 1801 CGGAAACGACAGAAATGTCAGCAATTAACCTGTGAAGCTTAAGATTCAGGGAACCTGAA 1860
DB 1798 CGGAAACGACAGAAATGTCAGCAATTAACCTGTGAAGCTTAAGATTCAGGGAACCTGAA 1857
QY 1861 GTGGAAGAAATGAGAACTGGAAGAGATCAATCTGAGCAGAGCTCAAAATTCAGGAG 1920
DB 1858 GTGGAAGAAATGAGAACTGGAAGAGATCAATCTGAGCAGAGCTCAAAATTCAGGAG 1917
QY 1921 CTTCAAGAGAACTGGAAGAGCTGTATAAGCCAGACAGAGGCAACCGACCTGTCGAG 1980
DB 1918 CTTCAAGAGAACTGGAAGAGCTGTATAAGCCAGACAGAGGCAACCGACCTGTCGAG 1977
QY 1981 AATATCCGCCAGGCAAGAGACGAGCCGAGAGGAGCTGGAAGAGCTGCAGAAACCGAGAG 2040
DB 1978 AATATCCGCCAGGCAAGAGACGAGCCGAGAGGAGCTGGAAGAGCTGCAGAAACCGAGAG 2037
QY 2041 GATTCTTGTGAAGGATTCAGAAAGAGTGTGAGAGCTGAGAGACGCGCATTTCTCTG 2100
DB 2038 GATTCTTGTGAAGGATTCAGAAAGAGTGTGAGAGCTGAGAGACGCGCATTTCTCTG 2097
QY 2101 GAGAACAGATTAAGAGACTAGAGACCATGAGCGTGAAGAAACAGACTAGAGATGAC 2160
DB 2098 GAGAACAGATTAAGAGACTAGAGACCATGAGCGTGAAGAAACAGACTAGAGATGAC 2157
QY 2161 ATCCAGACAAATTCCTCAACAGATTCAGCAGATGGCTGTATAAATTTCTGAGCTGAAAG 2220
DB 2158 ATCCAGACAAATTCCTCAACAGATTCAGCAGATGGCTGTATAAATTTCTGAGCTGAAAG 2217

QY 2221 AAACATCGGAGGCCCCAAGTCTCAGCCACAGCACTAGAAATGCACTTGAAACAGAAAGAG 2280
DB 2218 AAACATCGGAGGCCCCAAGTCTCAGCCACAGCACTAGAAATGCACTTGAAACAGAAAGAG 2277
QY 2281 CAGCACTATGAGAAAGATTTAAAGTGTGGAACATCATGATTAAGAAAGACTGGCTGAC 2340
DB 2278 CAGCACTATGAGAAAGATTTAAAGTGTGGAACATCATGATTAAGAAAGACTGGCTGAC 2337
QY 2341 AAGAGACATGAGAAACATGATGCAGAGACACGAGAGAGAGGCCCATGAGAGGCGAAA 2400
DB 2338 AAGAGACATGAGAAACATGATGCAGAGACACGAGAGAGAGGCCCATGAGAGGCGAAA 2397
QY 2401 ATTCTACGCCAACAAGAGGAGTATCATGCTATGGAATTCCTCAAGATCCATCCCTGAA 2460
DB 2398 ATTCTACGCCAACAAGAGGAGTATCATGCTATGGAATTCCTCAAGATCCCTGAA 2457
QY 2461 CAGAGATTGTGGAACCTGTGAAAGCCAAATTAACCTTGACAGAAATGACGCTTTTAAAC 2520
DB 2458 CAGAGATTGTGGAACCTGTGAAAGCCAAATTAACCTTGACAGAAATGACGCTTTTAAAC 2517
QY 2521 CAAAGAACATGAAAGGCCCAAGAAAGATGATTTCTGAACTCAGGCAACAGAAATTTTAC 2580
DB 2518 CAAAGAACATGAAAGGCCCAAGAAAGATGATTTCTGAACTCAGGCAACAGAAATTTTAC 2577
QY 2581 CTTGAGACACAGGCTGGAAAGTTGAGAGCCCAAGACGAAACTGAGAGGACGCTGGAG 2640
DB 2578 CTTGAGACACAGGCTGGAAAGTTGAGAGCCCAAGACGAAACTGAGAGGACGCTGGAG 2637
QY 2641 AAGATCAGCCACCAAGACCAAGTGAACAAGATCCGCTGTGGAATCTGAGACAAAGATTG 2700
DB 2638 AAGATCAGCCACCAAGACCAAGTGAACAAGATCCGCTGTGGAATCTGAGACAAAGATTG 2687
QY 2701 CGGAGGATCAGTCTAGAGCAGAGAGCAAGAACTGGAGCTCAAGCCGACTCACAGAG 2760
DB 2698 CGGAGGATCAGTCTAGAGCAGAGAGCAAGAACTGGAGCTCAAGCCGACTCACAGAG 2757
QY 2761 CTAACGCTTCCCTGACAGAGCCGAGTCAACAGTTGACAGCCCTGACGCTGCAAGGCG 2820
DB 2758 CTAACGCTTCCCTGACAGAGCCGAGTCAACAGTTGACAGCCCTGACGCTGCAAGGCG 2817
QY 2821 GCCCTGAGAGCCAGCTTCCGACAGGCGAAGACAGAGCTGGAAGAGACACAGCGAAGGCT 2880
DB 2818 GCCCTGAGAGCCAGCTTCCGACAGGCGAAGACAGAGCTGGAAGAGACACAGCGAAGGCT 2877
QY 2881 GAAAGAGAGATCCAGGACCTCAAGGCAATAGAGATGAATCCAGCCCAATTTGATGCT 2940
DB 2878 GAAAGAGAGATCCAGGACCTCAAGGCAATAGAGATGAATCCAGCCCAATTTGATGCT 2937
QY 2941 CTTCTGAACAGCTGTATCTGTATTCACAGACTGAGAGGCACTTAAACAGCTGACCGAG 3000
DB 2938 CTTCTGAACAGCTGTATCTGTATTCACAGACTGAGAGGCACTTAAACAGCTGACCGAG 2997
QY 3001 GACAACTCTGAATCTCAACCAAACTTTCTATTTTCCAAACAACTTCGATGAGGCTTCT 3060
DB 2998 GACAACTCTGAATCTCAACCAAACTTTCTATTTTCCAAACAACTTCGATGAGGCTTCT 3057
QY 3061 GGGGCCAAGCAGATTTGTACAACTGCGAAGTGAATGAGCACTCCGCGCGGAGATC 3120
DB 3058 GGGGCCAAGCAGATTTGTACAACTGCGAAGTGAATGAGCACTCCGCGCGGAGATC 3117
QY 3121 ACGGAACGAGAGATGCACTTACAGCCAGAAAGCAACGATGAGGCTTGAAGACGAG 3180
DB 3118 ACGGAACGAGAGATGCACTTACAGCCAGAAAGCAACGATGAGGCTTGAAGACGAG 3177
QY 3181 TGCACTATGCTGAGAGAAACAGTCTATGATTTTGAAGGCTTAAACGATGAGCTGTGAA 3240
DB 3178 TGCACTATGCTGAGAGAAACAGTCTATGATTTTGAAGGCTTAAACGATGAGCTGTGAA 3237
QY 3241 AAAGAGGCGAGTGGGAGGCTGGAGAGGCTCTGGGTGATGGAATCCCACTTTGAG 3300
DB 3238 AAAGAGGCGAGTGGGAGGCTGGAGAGGCTCTGGGTGATGGAATCCCACTTTGAG 3297
QY 3301 TGTGCGGTTGAGAGCTTGCAGAGAGATCTGGAACCGAGAAACAGAGCAAGGCGAGAGCC 3360

QY 5521 TACGAGAGCTAGCCGACAGACATCTCAAGTGAATGCTTACCTTGGCCTTGGCC 5580
DB 5518 TACGAGAGAGTACCGGACAGACAGATCTCAAGTGAATGCTTACCTTGGCCTTGGCC 5577
QY 5581 TACAGAGAACCCATCTGTTTGTGACCACTTCAACTCAAGTAATTTAGATCCAG 5640
DB 5578 TACAGAGAACCCATCTGTTTGTGACCACTTCAACTCAAGTAATTTAGATCCAG 5637
QY 5641 GCACGCTCTCAGACAGAGACCCCTGCGCAGCTACCTTGACATCCCGAACCCGCTAC 5700
DB 5638 GCACGCTCTCAGACAGAGACCCCTGCGCAGCTACCTTGACATCCCGAACCCGCTAC 5697
QY 5701 CTGGGCTCTGACATTTCTTCAAGAGCGATTACTTGGCGCTCTCATACAGATAATTA 5760
DB 5698 CTGGGCTCTGACATTTCTTCAAGAGCGATTACTTGGCGCTCTCATACAGATAATTA 5757
QY 5761 AGGATATTTGCTGCAAGAGAACTCTGTGAAGAGTCCGGCACTGAACCAACCGGGGC 5820
DB 5758 AGGATATTTGCTGCAAGAGAACTCTGTGAAGAGTCCGGCACTGAACCAACCGGGGC 5817
QY 5821 CCGTCCACCTCCCGACAGACCCCAACAAGAGAGCCCACTGACAGAGACATC 5880
DB 5818 CCGTCCACCTCCCGACAGACCCCAACAAGAGAGCCCACTGACAGAGACATC 5877
QY 5881 ACCAAGCGGCTGCTCTCAAGCCGCGCCGAAAGCCCAAGCCGCGAGAGCA 5940
DB 5878 ACCAAGCGGCTGCTCTCAAGCCGCGCCGAAAGCCCAAGCCGCGAGAGCA 5937
QY 5941 AGACACACCCCAAGCTACCGCGAGAGGCGGACCGAGTGGCGAGGACAACTCTCTGGC 6000
DB 5938 AGACACACCCCAAGCTACCGCGAGAGGCGGACCGAGTGGCGAGGACAACTCTCTGGC 5997
QY 6001 CGCCCTCTGAGAGAGAGAAATCCCGCGCGGATCTCAGCAAGCGAGAGAGCGGTCC 6060
DB 5998 CGCCCTCTGAGAGAGAGAAATCCCGCGCGGATCTCAGCAAGCGAGAGAGCGGTCC 6057
QY 6061 CCGGACAGGCTGTTTGAAGACAGACAGAGAGGCGCGCTGCTGCGGAGCGGTGAGAAC 6120
DB 6058 CCGGAGAGGCTGTTTGAAGACAGACAGAGAGGCGCGCTGCTGCGGAGCGGTGAGAAC 6117
QY 6121 CCGCTGTCCAGGTGAACAAGTCTGGAGCAATC 6155
DB 6118 CCGCTGTCCAGGTGAACAAGTCTGGAGCAATC 6152

RESULT 9
US-10-028-946-3
Sequence 3, Application US/10028946
Publication No. US20020123622A1
GENERAL INFORMATION:

APPLICANT: Yu, Xuanhuan
APPLICANT: Miranda, Maricar
APPLICANT: Fridlie, Carl Johan
TITLE OF INVENTION: No. US20020123622A1 Human Kinases and Polynucleotides Encoding
FILE REFERENCE: LEX-0289-USA
CURRENT APPLICATION NUMBER: US/10/028,946
CURRENT FILING DATE: 2001-12-20
PRIOR APPLICATION NUMBER: US 60/258,335
PRIOR FILING DATE: 2000-12-27
NUMBER OF SEQ ID NOS: 4
SOFTWARE: FaastSeq for Windows Version 4.0
SEQ ID NO 3
LENGTH: 5877
TYPE: DNA
ORGANISM: homo sapiens
US-10-028-946-3

Query Match 94.7%; Score 5837; DB 13; Length 5877;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 5837; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGTTGAAGTTCAAAATATGAGCGCGAATCTTTGGATGCTGTGCTGTAACCATTT 60

DB 1 ATGTTGAAGTTCAAAATATGAGCGCGAATCTTTGGATGCTGTGCTGTAACCATTT 60
QY 61 GCCAGCGGAGCTCCAGAGCTGAATCTGTTCTTCCAGAGGAAACCAACCTTTATGACTCAA 120
DB 61 GCCAGCGGAGCTCCAGAGCTGAATCTGTTCTTCCAGAGGAAACCAACCTTTATGACTCAA 120
QY 121 CAGCAGATGCTCTCTTCTTCCGAGAGGAGATTAAGATGCTCTTGTCTCTTTGAA 180
DB 121 CAGCAGATGCTCTCTTCTTCCGAGAGGAGATTAAGATGCTCTTGTCTCTTTGAA 180
QY 181 GAATGAGTCAAGCTGCTCTGATGAAGATTAAGCAAGTGAACAATTTGTCGGAAATAT 240
DB 181 GAATGAGTCAAGCTGCTCTGATGAAGATTAAGCAAGTGAACAATTTGTCGGAAATAT 240
QY 241 TCCGACACCATAGCTAGTTACAGAGCTCCAGCTTCCGCAAGAGACTTCCAGATCAGA 300
DB 241 TCCGACACCATAGCTAGTTACAGAGCTCCAGCTTCCGCAAGAGACTTCCAGATCAGA 300
QY 301 AGTCTTGAAGTGTGTGCTCACTTGTCTGAAGTGCAGTGTGAAGAGAAAGCAACCGG 360
DB 301 AGTCTTGAAGTGTGTGCTCACTTGTCTGAAGTGCAGTGTGAAGAGAAAGCAACCGG 360
QY 361 GACATCTATGCTATGAAGTGAAGAGAGAGGCTTTATGGCCAGAGAGAGCTTTCA 420
DB 361 GACATCTATGCTATGAAGTGAAGAGAGAGGCTTTATGGCCAGAGAGAGCTTTCA 420
QY 421 TTTTGAAGAGAGGAGCAATATATCTGACAGAGCAAGCCGCTGATCCCAATTA 480
DB 421 TTTTGAAGAGAGGAGCAATATATCTGACAGAGCAAGCCGCTGATCCCAATTA 480
QY 481 CAGTATGCTTTCAAGACAAATCACTTTATCTGTGATGGAATATCAGCTGAGAGG 540
DB 481 CAGTATGCTTTCAAGACAAATCACTTTATCTGTGATGGAATATCAGCTGAGAGG 540
QY 541 GACTTCTGCTCACTTTGTAATGATATGAGAGACAGTGAATGAATGAATGATGATG 600
DB 541 GACTTCTGCTCACTTTGTAATGATATGAGAGACAGTGAATGAATGAATGATGATG 600
QY 601 TACCTAGCTGAGCTGATTTTGGCTGTTCAAGAGCTCATCTGATGAGATATGATGATCA 660
DB 601 TACCTAGCTGAGCTGATTTTGGCTGTTCAAGAGCTCATCTGATGAGATATGATGATCA 660
QY 661 GACATCAAGCTGAGAAATCTCTGTTGACCGACAGACACATCAAGCTGATGATTT 720
DB 661 GACATCAAGCTGAGAAATCTCTGTTGACCGACAGACACATCAAGCTGATGATTT 720
QY 721 GGAATGCTGCGGAAATGAATTCAAACAAGATGTAATGCCAACTCCGATTTGGAGC 780
DB 721 GGAATGCTGCGGAAATGAATTCAAACAAGATGTAATGCCAACTCCGATTTGGAGC 780
QY 781 CCAATTAATGAGCTCTGTAAGTCTGATGATGAAGGAGATGGAAGGACACTAC 840
DB 781 CCAATTAATGAGCTCTGTAAGTCTGATGATGAAGGAGATGGAAGGACACTAC 840
QY 841 GGCCTGAGCTGACTGCTGATGAGAGGAGGCTGATGCTGATGAGATGATTTAGAGAGA 900
DB 841 GGCCTGAGCTGACTGCTGATGAGAGGAGGCTGATGCTGATGAGATGATTTAGAGAGA 900
QY 901 TCCCTCTGAGAGGAACTCTGCGAAGACTTCAATTAATTAATGATTTCCAGCGG 960
DB 901 TCCCTCTGAGAGGAACTCTGCGAAGACTTCAATTAATTAATGATTTCCAGCGG 960
QY 961 TTTTGAATTTCCAGATGACCCCAAGAGAGAGAGCTTTTATGATGATTTCAAGC 1020
DB 961 TTTTGAATTTCCAGATGACCCCAAGAGAGAGAGCTTTTATGATGATTTCAAGC 1020
QY 1021 TTGTTGTGCGGACAGAAAGAGAGACTGAAGTCTTGTCTGCACTCTTCTTC 1080
DB 1021 TTGTTGTGCGGACAGAAAGAGAGACTGAAGTCTTGTCTGCACTCTTCTTC 1080
QY 1081 TCTAAATGACTGGAACAATCTGTAATCTCTCTCCCTGTTGCTTCCACCTCAAG 1140

Db 1081 TCTAAATGACTGGAACAATTCTGTACTCTCCCTCCCTGTTCCACCCTCAAG 1140
Qy 1141 TCTGACATGACCTCTCCAAATTTTGATGAACGAGAAAGATTGTGGGTTTCATCTCT 1200
Db 1141 TCTGACATGACCTCTCCAAATTTTGATGAACGAGAAAGATTGTGGGTTTCATCTCT 1200
Qy 1201 CCGTGCAGCTGAGCCCTCAGGCTCTCGGGTGAAGAACTCGCGTTTGTGGGGTTTTCG 1260
Db 1201 CCGTGCAGCTGAGCCCTCAGGCTCTCGGGTGAAGAACTCGCGTTTGTGGGGTTTTCG 1260
Qy 1261 TAGACGAGGCACTGGGATTTCTTGATAGTCTGAGTCTGTGTGTGCGGCTTGAATCC 1320
Db 1261 TAGACGAGGCACTGGGATTTCTTGATAGTCTGTGTGTGCGGCTTGAATCC 1320
Qy 1321 CTTGCCAAGCTAGCTCATGGAAGAACTTCTCATCAAAAGCAAAAGCTACAGAC 1380
Db 1321 CTTGCCAAGCTAGCTCATGGAAGAACTTCTCATCAAAAGCAAAAGCTACAGAC 1380
Qy 1381 TCTCAGGACAAGTGTCAAGATGAGGCAAGAAATGACCCGTTACATCGGAGATGTCA 1440
Db 1381 TCTCAGGACAAGTGTCAAGATGAGGCAAGAAATGACCCGTTACATCGGAGATGTCA 1440
Qy 1441 GAGGTGAGGCTGTGCTTAAGTCAAGAGAGTGAAGGCTTGAAGCTTGAAGCTCAGAGA 1500
Db 1441 GAGGTGAGGCTGTGCTTAAGTCAAGAGAGTGAAGGCTTGAAGCTTGAAGCTCAGAGA 1500
Qy 1501 TCCCTCTCTGAGCAGACCTTGTCTACCTACATCAAGATGCAAGTCTTAAAGCAAGT 1560
Db 1501 TCCCTCTCTGAGCAGACCTTGTCTACCTACATCAAGATGCAAGTCTTAAAGCAAGT 1560
Qy 1561 TTGGAGCAGCAGGATGAGGATGCTCCAGAGAGATGCAAGCACTGACGCTTTCAT 1620
Db 1561 TTGGAGCAGCAGGATGAGGATGCTCCAGAGAGATGCAAGCACTGACGCTTTCAT 1620
Qy 1621 GATATCAGAGCAGAGCCCGGAAGCTCAAGAAATCAAAAGCAGAGTACAGGCTCAA 1680
Db 1621 GATATCAGAGCAGAGCCCGGAAGCTCAAGAAATCAAAAGCAGAGTACAGGCTCAA 1680
Qy 1681 GTGGAAAGAAATGAGTTGATGATGAATCAATGTTGAAAGAGATCTTGTCTCAGCAAGAA 1740
Db 1681 GTGGAAAGAAATGAGTTGATGATGAATCAATGTTGAAAGAGATCTTGTCTCAGCAAGAA 1740
Qy 1741 CGGAGTATCTCTACGAATCTGAGCTGAGAGAGTCTCGGTTGCTCTGAAAGATTCAAG 1800
Db 1741 CGGAGTATCTCTACGAATCTGAGCTGAGAGAGTCTCGGTTGCTCTGAAAGATTCAAG 1800
Qy 1801 CGGAAAGCAGCAAGATGTGAGCATAACTGTTGAAGCTTAAGGATCAAGGAAAGCTGAA 1860
Db 1801 CGGAAAGCAGCAAGATGTGAGCATAACTGTTGAAGCTTAAGGATCAAGGAAAGCTGAA 1860
Qy 1861 GTGGGAGATATGTCGAACTGAGAAAGATCAATGCTGAGCAGAGCTCAAAATTCAGAG 1920
Db 1861 GTGGGAGATATGTCGAACTGAGAAAGATCAATGCTGAGCAGAGCTCAAAATTCAGAG 1920
Qy 1921 CTCCAAAGAAATCTGAGAAAGCTGTTAAAGCCAGACCGAGGCTGAGAACTGCAAG 1980
Db 1921 CTCCAAAGAAATCTGAGAAAGCTGTTAAAGCCAGACCGAGGCTGAGAACTGCAAG 1980
Qy 1981 AATATCCGCAAGCAAGAGCGAGCGAGAGGAGCTGAGAAAGCTGAGAAACGAGAG 2040
Db 1981 AATATCCGCAAGCAAGAGCGAGCGAGAGGAGCTGAGAAAGCTGAGAAACGAGAG 2040
Qy 2041 GATTCTTCTGAAGGCACTAGAAAGAAAGCTGTGGAAGCTGAGAAAGCCGCAATTCCTG 2100
Db 2041 GATTCTTCTGAAGGCACTAGAAAGAAAGCTGTGGAAGCTGAGAAAGCCGCAATTCCTG 2100
Qy 2101 GAGAACAAAGTAAAGAGACTAGAGACCATGAGAGCTGAGAAAGCACTGAAAGGTGAC 2160
Db 2101 GAGAACAAAGTAAAGAGACTAGAGACCATGAGAGCTGAGAAAGCACTGAAAGGTGAC 2160
Qy 2161 ATCCAGACAAATCCCAACAGATCCAGCAGATGGCTGATAAATTCGAGGCTCGAAGAG 2220
Db 2161 ATCCAGACAAATCCCAACAGATCCAGCAGATGGCTGATAAATTCGAGGCTCGAAGAG 2220

Qy 2221 AAACATCGGAGGCCCAAGTCTCAGCCAGCACCTAGAAAGTGAACCTGAAGCAGAAAGAG 2280
Db 2221 AAACATCGGAGGCCCAAGTCTCAGCCAGCACCTAGAAAGTGAACCTGAAGCAGAAAGAG 2280
Qy 2281 CAGCACTATGAGAAAGAAATTAAGTTGAGCAATCAAGTAAAGAAAGCTGGCTGAC 2340
Db 2281 CAGCACTATGAGAAAGAAATTAAGTTGAGCAATCAAGTAAAGAAAGCTGGCTGAC 2340
Qy 2341 AAGGAGCACTGAGAAACATGATCAGAGACAAGAGAGAGGCCCATGAGAAAGGCAAA 2400
Db 2341 AAGGAGCACTGAGAAACATGATCAGAGACAAGAGAGAGGCCCATGAGAAAGGCAAA 2400
Qy 2401 ATTCTAGCGAACAAGAGGATGATCATGATGATGATTCAGAGATTCAGATTCCTGAA 2460
Db 2401 ATTCTAGCGAACAAGAGGATGATCATGATGATGATTCAGAGATTCAGATTCCTGAA 2460
Qy 2461 CAGAGATTTGTGAATCTGTGAGAGCAATTAACCTGAGCAAAATGACGCTTTTACC 2520
Db 2461 CAGAGATTTGTGAATCTGTGAGAGCAATTAACCTGAGCAAAATGACGCTTTTACC 2520
Qy 2521 CAAAGGAACATGAAGGCCCAAGAAAGATGATTTCTGAATCTCAGGCAACAGAAATTTAC 2580
Db 2521 CAAAGGAACATGAAGGCCCAAGAAAGATGATTTCTGAATCTCAGGCAACAGAAATTTAC 2580
Qy 2581 CTGGAGACAAGGCTGGAAAGTTGAGAGCCAGAACCGAAACTGAGAGAGCCTGAG 2640
Db 2581 CTGGAGACAAGGCTGGAAAGTTGAGAGCCAGAACCGAAACTGAGAGAGCCTGAG 2640
Qy 2641 AAGATAGCCACCAAGAACCAAGTGAAGAAATCGGCTCTGAAATGAGACAAATTTG 2700
Db 2641 AAGATAGCCACCAAGAACCAAGTGAAGAAATCGGCTCTGAAATGAGACAAATTTG 2700
Qy 2701 CCGGAGGTCACTGAGAGCAGAGAGCAGAACTGAACTCAAGGCCAGCTCAGAGAG 2760
Db 2701 CCGGAGGTCACTGAGAGCAGAGAGCAGAACTGAACTCAAGGCCAGCTCAGAGAG 2760
Qy 2761 CTACAGCTCTCCCTGAGAGGCGGAGTCAAGTTGACAGCCCTGAGGCTCAGCGGCG 2820
Db 2761 CTACAGCTCTCCCTGAGAGGCGGAGTCAAGTTGACAGCCCTGAGGCTCAGCGGCG 2820
Qy 2821 GCCCTGAGAGGCAAGCTTGGCAGAGGAGAGAGAGCTGGAAGAGCCACAGAAAGCT 2880
Db 2821 GCCCTGAGAGGCAAGCTTGGCAGAGGAGAGAGAGCTGGAAGAGCCACAGAAAGCT 2880
Qy 2881 GAAAGAGATCAGGCACTCAGGCACTAGAGATGAATCAAGGCCAAATTTGATGCT 2940
Db 2881 GAAAGAGATCAGGCACTCAGGCACTAGAGATGAATCAAGGCCAAATTTGATGCT 2940
Qy 2941 GTTCGTAACAGCTGTATCTGTAACTCAGACTGAGAGAGCAGCTAAACAGCTGACGAG 3000
Db 2941 GTTCGTAACAGCTGTATCTGTAACTCAGACTGAGAGAGCAGCTAAACAGCTGACGAG 3000
Qy 3001 GACAAAGCTGAATCTCAACCAAACTTTCTCTTCCAAACAATCTGATGAGGTTCT 3060
Db 3001 GACAAAGCTGAATCTCAACCAAACTTTCTCTTCCAAACAATCTGATGAGGTTCT 3060
Qy 3061 GCGCCCAAGCAGAGATTTGTAACACTGCGAAGTGAAGTGAACATCTCCGCGGAGATC 3120
Db 3061 GCGCCCAAGCAGAGATTTGTAACACTGCGAAGTGAAGTGAACATCTCCGCGGAGATC 3120
Qy 3121 ACGGAACGAGATCAGCTTACAGCCAGCAAGCAAGCAAGATGAGGCTTGAAGACAG 3180
Db 3121 ACGGAACGAGATCAGCTTACAGCCAGCAAGCAAGCAAGATGAGGCTTGAAGACAG 3180
Qy 3181 TGCAACATGCTGAGAGAAACAGGTCACTGATTTGAGAGCCCTTAAAGATGAGCTGTAGAA 3240
Db 3181 TGCAACATGCTGAGAGAAACAGGTCACTGATTTGAGAGCCCTTAAAGATGAGCTGTAGAA 3240
Qy 3241 AAAGAGCGCAGTGGAGGCTCGAGAGAGCGTCTGGGATGAGAAATCCAGTTTGAAG 3300
Db 3241 AAAGAGCGCAGTGGAGGCTCGAGAGAGCGTCTGGGATGAGAAATCCAGTTTGAAG 3300

QY 3301 TGTGGTTGAGAGCTGCAAGAGATGCTGSAACCCGAGAAACAGAGCAAGGCGAGAGCC 3360
DB 3301 TGTGGTTGAGAGCTGCAAGAGATGCTGSAACCCGAGAAACAGAGCAAGGCGAGAGCC 3360
QY 3361 GATCAGCGATCAACCGAGTCTCCGCAAGTGGAGCTGGCAGTGAAGAGACAAAGCT 3420
DB 3361 GATCAGCGATCAACCGAGTCTCCGCAAGTGGAGCTGGCAGTGAAGAGACAAAGCT 3420
QY 3421 GAGATTCTGCTCTGCAAGAGGCTCTCAAGAGAGCAAGAGTGAAGGCCGAGAGCTCTCT 3480
DB 3421 GAGATTCTGCTCTGCAAGAGGCTCTCAAGAGAGCAAGAGTGAAGGCCGAGAGCTCTCT 3480
QY 3481 GACAAGCTCAATGACTCTGAGAGAGAGAGATGCTATGCTTGAATGAATGCCGAAGCTTA 3540
DB 3481 GACAAGCTCAATGACTCTGAGAGAGAGAGATGCTATGCTTGAATGAATGCCGAAGCTTA 3540
QY 3541 CAGAGAGAGCTGAGAGCTGAACGAGAGCTCAACAGAGGCTTCTGAGAGAGCAAGCCAA 3600
DB 3541 CAGAGAGAGCTGAGAGCTGAACGAGAGCTCAACAGAGGCTTCTGAGAGAGCAAGCCAA 3600
QY 3601 TTACAGCAGAGATGAGACTGCAAGAAATCAACATTTCCGTCTGACTCAAGAGACTGCAA 3660
DB 3601 TTACAGCAGAGATGAGACTGCAAGAAATCAACATTTCCGTCTGACTCAAGAGACTGCAA 3660
QY 3661 GAAAGCTTAAGATCGAGCTGATCTATCTGAAGACAGAAAGAAAGTGAAGTATCAAGCTG 3720
DB 3661 GAAAGCTTAAGATCGAGCTGATCTATCTGAAGACAGAAAGAAAGTGAAGTATCAAGCTG 3720
QY 3721 GAAAACTTCAAGTCTCTATCTTCAATGAAGAGTGAAGTGAAGAGCATATTTCTCAA 3780
DB 3721 GAAAACTTCAAGTCTCTATCTTCAATGAAGAGTGAAGTGAAGAGCATATTTCTCAA 3780
QY 3781 CAAACCAACTCATTTGATTTTCTGCAAGCCAAATGAGACCACTGCTAAGAAAG 3840
DB 3781 CAAACCAACTCATTTGATTTTCTGCAAGCCAAATGAGACCACTGCTAAGAAAG 3840
QY 3841 GTTCTCTGAGATCAATAGAGTGAAGCTGAGCTGAGAGAGAAAGTCTGCTGTGCA 3900
DB 3841 GTTCTCTGAGATCAATAGAGTGAAGCTGAGCTGAGAGAGAAAGTCTGCTGTGCA 3900
QY 3901 GAGCTAAGAGAGCCCTTCAAGAGACCCGATCGAGCTCCGCTCCGCGAGAGAGAGCT 3960
DB 3901 GAGCTAAGAGAGCCCTTCAAGAGACCCGATCGAGCTCCGCTCCGCGAGAGAGAGCT 3960
QY 3961 GCCCAGCGAAAG 4020
DB 3961 GCCCAGCGAAAG 4020
QY 4021 ATGCGCATGTCGCGCATCTGTCGAGTGCAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4080
DB 4021 ATGCGCATGTCGCGCATCTGTCGAGTGCAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4080
QY 4081 GCCCGCGCATCTGAG 4140
DB 4081 GCCCGCGCATCTGAG 4140
QY 4141 AAGGAG 4200
DB 4141 AAGGAG 4200
QY 4201 ACAAG 4260
DB 4201 ACAAG 4260
QY 4261 GAATGTCAAGT 4320
DB 4261 GAATGTCAAGT 4320
QY 4321 CTTGCTGAATATGCAACACTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4380
DB 4321 CTTGCTGAATATGCAACACTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4380
QY 4381 GGTCTCCAGACCAAG 4440

DB 4381 GGTCTCCAGACCAAG 4440
QY 4441 AGGAATTAACAAAG 4500
DB 4441 AGGAATTAACAAAG 4500
QY 4501 AAGTCTCAATTTATGACAAATGAAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4560
DB 4501 AAGTCTCAATTTATGACAAATGAAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4560
QY 4561 GAGTGTGCTTCCGAGCGGAG 4620
DB 4561 GAGTGTGCTTCCGAGCGGAG 4620
QY 4621 GCAATTAACAG 4680
DB 4621 GCAATTAACAG 4680
QY 4681 ACCTGTGAGCGCGGAG 4740
DB 4681 ACCTGTGAGCGCGGAG 4740
QY 4741 TGGGTACCGGCTTGAATGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4800
DB 4741 TGGGTACCGGCTTGAATGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4800
QY 4801 GCTGATGCTAACTGCTGTGAGAACTCCCTGCTGAACTGGAAGAGTGAACCGTCTAGAC 4860
DB 4801 GCTGATGCTAACTGCTGTGAGAACTCCCTGCTGAACTGGAAGAGTGAACCGTCTAGAC 4860
QY 4861 ATGACTGACAGCTGAGCTTCAAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4920
DB 4861 ATGACTGACAGCTGAGCTTCAAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4920
QY 4921 TAGGCTGAGATGCTTGAAGAACTCCCTGAACCATCTCCAGAGAGAGAGAGAGAGAG 4980
DB 4921 TAGGCTGAGATGCTTGAAGAACTCCCTGAACCATCTCCAGAGAGAGAGAGAGAGAG 4980
QY 4981 CAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 5040
DB 4981 CAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 5040
QY 5041 CTGTGCTTGTGAG 5100
DB 5041 CTGTGCTTGTGAG 5100
QY 5101 CAGCCGAGATCTCAACCAATTTTGAAGCTGTCAAGAGAGAGAGAGAGAGAGAGAGAG 5160
DB 5101 CAGCCGAGATCTCAACCAATTTTGAAGCTGTCAAGAGAGAGAGAGAGAGAGAGAGAG 5160
QY 5161 GGCAGATTGAG 5220
DB 5161 GGCAGATTGAG 5220
QY 5221 CGCTACAG 5280
DB 5221 CGCTACAG 5280
QY 5281 TGAGCTGTATCACTTCAACCAATTAAGTGTCTATTTGAGAGAGAGAGAGAGAGAGAG 5340
DB 5281 TGAGCTGTATCACTTCAACCAATTAAGTGTCTATTTGAGAGAGAGAGAGAGAGAGAG 5340
QY 5341 ATGAGCATGAAG 5400
DB 5341 ATGAGCATGAAG 5400
QY 5401 CTTGCTGTGTTGCGGCTCTTCAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5460
DB 5401 CTTGCTGTGTTGCGGCTCTTCAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5460
QY 5461 GCAGGAG 5520

Db 5461 GCAGGCGAGGAGAGTACTGTGTGTTTCCACGAATTTGGAGTGTCTGATTTCT 5520
Qy 5521 TACGGAAGAGCTTACCGGACAGAGATCTCAATGAGAGTGGCTTACCTTTGGCCCTTTGCC 5580
Db 5521 TACGGAAGAGCTTACCGGACAGAGATCTCAATGAGAGTGGCTTACCTTTGGCCCTTTGCC 5580
Qy 5581 TACGGAAGAGCTTACCGGACAGAGATCTCAATGAGAGTGGCTTACCTTTGGCCCTTTGCC 5580
Db 5581 TACGGAAGAGCTTACCGGACAGAGATCTCAATGAGAGTGGCTTACCTTTGGCCCTTTGCC 5580
Qy 5641 GCAGGCTCTCAGCAGGAGACCCCTGCGGAGCGTACCTGACATCCGGAACCCGCGCTAC 5700
Db 5641 GCAGGCTCTCAGCAGGAGACCCCTGCGGAGCGTACCTGACATCCGGAACCCGCGCTAC 5700
Qy 5701 CTGGGCCCCCTGCAATTTCTCTCAGAGCGATTTACTTGGCGCTCTCATACAGATTAATTA 5760
Db 5701 CTGGGCCCCCTGCAATTTCTCTCAGAGCGATTTACTTGGCGCTCTCATACAGATTAATTA 5760
Qy 5761 AGGGTCATTTGCTGCAAGGAGAACTGTGAAGAGTCCGCACTGAACCAACCGGGGC 5820
Db 5761 AGGGTCATTTGCTGCAAGGAGAACTGTGAAGAGTCCGCACTGAACCAACCGGGGC 5820
Qy 5821 CCGTCCACCTCCGCGAG 5837
Db 5821 CCGTCCACCTCCGCGAG 5837

RESULT 10
US-10-791-666-3
/ Sequence 3, Application US/10791666
/ Publication No. US20040209297A1
/ GENERAL INFORMATION:
/ APPLICANT: Yu, Xuanhuan
/ APPLICANT: Miranda, Maricar
/ APPLICANT: Fiddle, Carl Johan
/ TITLE OF INVENTION: Novel Human Kinases and Polynucleotides Encoding the Same
/ FILE REFERENCE: LEX-0289-USA
/ CURRENT APPLICATION NUMBER: US/10/791,666
/ CURRENT FILING DATE: 2004-03-02
/ PRIOR APPLICATION NUMBER: US/10/028,946
/ PRIOR FILING DATE: 2001-12-20
/ PRIOR APPLICATION NUMBER: US 60/258,335
/ PRIOR FILING DATE: 2000-12-27
/ NUMBER OF SEQ ID NOS: 4
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 3
/ LENGTH: 5877
/ TYPE: DNA
/ ORGANISM: homo sapiens
US-10-791-666-3

Query Match 94.7%; Score 5837; DB 18; Length 5877;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 5837; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AGTTTAAAGTTCAAATATGAGAGCGGGAATCTTTGATGCTGCTGCTGAACCCATT 60
Db 1 AGTTTAAAGTTCAAATATGAGAGCGGGAATCTTTGATGCTGCTGCTGAACCCATT 60
Qy 61 GCCAGCGGAGCCCTCCAGGCTGAATCTGTTCTTCAGAGGGAACACCCCTTTAAGACTCA 120
Db 61 GCCAGCGGAGCCCTCCAGGCTGAATCTGTTCTTCAGAGGGAACACCCCTTTAAGACTCA 120
Qy 121 CAGCAGATGTCTCTCTTTCCGAGAAAGGATATTAGATGCCCTCTTTGTTCTTTGAA 180
Db 121 CAGCAGATGTCTCTCTTTCCGAGAAAGGATATTAGATGCCCTCTTTGTTCTTTGAA 180
Qy 181 GAATGAGTCAAGCTGCTCTGATGAAGATTAAAGCAAGTGAACCTTTGTCGGGAAGTAT 240
Db 181 GAATGAGTCAAGCTGCTCTGATGAAGATTAAAGCAAGTGAACCTTTGTCGGGAAGTAT 240
Qy 241 TCCGACACATAGCTGAGTTACAGAGCTCCAGCCTTCGCAAGGACTTTCGAAGTCAAA 300
Db 241 TCCGACACATAGCTGAGTTACAGAGCTCCAGCCTTCGCAAGGACTTTCGAAGTCAAA 300

Db 241 TCCGACACATAGCTGAGTTACAGAGCTCCAGCCTTCGCAAGGACTTTCGAAGTCAAA 300
Qy 301 AGTCTTAGAGTGTGTGCTCACTTTGCTGAAGTGCAGGTGTGTAAGAGAAAGCAACCGGG 360
Db 301 AGTCTTAGAGTGTGTGCTCACTTTGCTGAAGTGCAGGTGTGTAAGAGAAAGCAACCGGG 360
Qy 361 GACATCTATGCTATGAAGAGTGAAGAAAGAGGCTTTATTTGGCCCGAGAGCAGGTTTCA 420
Db 361 GACATCTATGCTATGAAGAGTGAAGAAAGAGGCTTTATTTGGCCCGAGAGCAGGTTTCA 420
Qy 421 TTTTGTAGGAAAGGCGGAACATATTATCTGAAGCAAGGCCGAGGATCCCAATTTA 480
Db 421 TTTTGTAGGAAAGGCGGAACATATTATCTGAAGCAAGGCCGAGGATCCCAATTTA 480
Qy 481 CAGTATGCCCTTTAGGACAAATAATCACTTTATCTGTGTCATGTAATATACGCTGAGGG 540
Db 481 CAGTATGCCCTTTAGGACAAATAATCACTTTATCTGTGTCATGTAATATACGCTGAGGG 540
Qy 541 GACTTGCTGTCACTTTTGAATTAATATGAGACCAAGTTAGATGAAAACCTGATACGTTT 600
Db 541 GACTTGCTGTCACTTTTGAATTAATATGAGACCAAGTTAGATGAAAACCTGATACGTTT 600
Qy 541 GACTTGCTGTCACTTTTGAATTAATATGAGACCAAGTTAGATGAAAACCTGATACGTTT 600
Db 541 GACTTGCTGTCACTTTTGAATTAATATGAGACCAAGTTAGATGAAAACCTGATACGTTT 600
Qy 601 TACCTAGCTGAGCTGATTTTGGCGTTTCAAGGGTTCATGTAATGGAATACGTGACATGA 660
Db 601 TACCTAGCTGAGCTGATTTTGGCGTTTCAAGGGTTCATGTAATGGAATACGTGACATGA 660
Qy 661 GACATCAAGCCTTAGAATCTCTGTGACCGGACAGACACATCAAGCTGTGTGATTTT 720
Db 661 GACATCAAGCCTTAGAATCTCTGTGACCGGACAGACACATCAAGCTGTGTGATTTT 720
Qy 721 GGATCTGCGCGGAAATGAATTTAAACAGATGTGTGAATGCCAACTCCGATTTGGAGCC 780
Db 721 GGATCTGCGCGGAAATGAATTTAAACAGATGTGTGAATGCCAACTCCGATTTGGAGCC 780
Qy 781 CCAGATTAACATGAGCTCTGTAAGTGTGACGTGATGTAAGGAGGATGAAAGGACACTAC 840
Db 781 CCAGATTAACATGAGCTCTGTAAGTGTGACGTGATGTAAGGAGGATGAAAGGACACTAC 840
Qy 841 GGCCTGAGCTGTGACTGTGTGTGCTGAGTGGGCGTGAATGCTATGAGATGATTTATGGGAGA 900
Db 841 GGCCTGAGCTGTGACTGTGTGTGCTGAGTGGGCGTGAATGCTATGAGATGATTTATGGGAGA 900
Qy 901 TCCGCTTTGCGAGAGGAACTCTGCGCAACCTTCAATPACTTATGAATTTCCAGCGG 960
Db 901 TCCGCTTTGCGAGAGGAACTCTGCGCAACCTTCAATPACTTATGAATTTCCAGCGG 960
Qy 961 TTTTGAATTTCCAGATGATACCCCAAGAGAGAGGAGGACTTCTTGAATGATTTCAAGC 1020
Db 961 TTTTGAATTTCCAGATGATACCCCAAGAGAGGAGGAGGACTTCTTGAATGATTTCAAGC 1020
Qy 1021 TTGTTGTGCGGCGCAGAAAGAGAGACTGAAGTCTTTGCTGCGCATCTTCTTCTTC 1080
Db 1021 TTGTTGTGCGGCGCAGAAAGAGAGACTGAAGTCTTTGCTGCGCATCTTCTTCTTC 1080
Qy 1081 TCTAAATTTGACTGGAACAACTTGTAACTCTCTCCCTCCCTTGCCTCCCAAG 1140
Db 1081 TCTAAATTTGACTGGAACAACTTGTAACTCTCTCCCTCCCTTGCCTCCCAAG 1140
Qy 1141 TCTGAAGATGACACCTCCCAATTTGATGAACAAGAAATTCGAGGCTTCACTCTCT 1200
Db 1141 TCTGAAGATGACACCTCCCAATTTGATGAACAAGAAATTCGAGGCTTCACTCTCT 1200
Qy 1201 CCGTGCAGCTGAGCCCTCAGGCTTCTGAGGTGAAGAACTCCGTTGTGAGGCTTTTGC 1260
Db 1201 CCGTGCAGCTGAGCCCTCAGGCTTCTGAGGTGAAGAACTCCGTTGTGAGGCTTTTGC 1260
Qy 1261 TACAGCAAGGCACTGAGGATCTTGTGTGATCTGATGCTGTGTGTGCGGCTCTGAGCTCC 1320
Db 1261 TACAGCAAGGCACTGAGGATCTTGTGTGATCTGATGCTGTGTGTGCGGCTCTGAGCTCC 1320
Qy 1321 CCGTGCAGCACTAGCTCCATGGAAGAAAGAACTTCTCATCAAAAGCAAGAGCTTACAGAC 1380
Db 1321 CCGTGCAGCACTAGCTCCATGGAAGAAAGAACTTCTCATCAAAAGCAAGAGCTTACAGAC 1380

QY 1381 TCTCAGGACAAAGTGTCAACAGATGAGACAGAAATGACCCCGGTTACATCCGAGAGTGTCA 1440
DB 1381 TCTCAGGACAAAGTGTCAACAGATGAGACAGAAATGACCCCGGTTACATCCGAGAGTGTCA 1440
QY 1441 GAGGTGAGGCTGTGCTTAACTGACAGAGAGAGTGTGAGAGGCTGTGAGACTCAGAGA 1500
DB 1441 GAGGTGAGGCTGTGCTTAACTGACAGAGAGAGTGTGAGAGGCTGTGAGACTCAGAGA 1500
QY 1501 TCCCTCCGAGACAGGACCTTGTACCTACATCAAGAAATGACAGATGAGCTTAAAGCGAAGT 1560
DB 1501 TCCCTCCGAGACAGGACCTTGTACCTACATCAAGAAATGACAGATGAGCTTAAAGCGAAGT 1560
QY 1561 TTGGAGCAAGCAGGATGAGAGTGTCCAGAGAGATGACAAAGACCTGACGCTTCCAT 1620
DB 1561 TTGGAGCAAGCAGGATGAGAGTGTCCAGAGAGATGACAAAGACCTGACGCTTCCAT 1620
QY 1621 GATATCAGAGACAGACCCGAGACCTCCAGAAATCAAGAGCAGAGTACAGGCTCAA 1680
DB 1621 GATATCAGAGACAGACCCGAGACCTCCAGAAATCAAGAGCAGAGTACAGGCTCAA 1680
QY 1681 GTGAGAGAAATGAGGTTGATGATCAAGTTGAGAGGATCTTGTCTCAGAGAGAA 1740
DB 1681 GTGAGAGAAATGAGGTTGATGATCAAGTTGAGAGGATCTTGTCTCAGAGAGAA 1740
QY 1741 CGGAGTATCTCTACGATCTGAGCTGAGAGAGTCTCGGCTTGTGCTGAGAAATTCAG 1800
DB 1741 CGGAGTATCTCTACGATCTGAGCTGAGAGAGTCTCGGCTTGTGCTGAGAAATTCAG 1800
QY 1801 CGGAAAGCAGAGATGTACGATTAACCTGTTGAAGCTTAAGATCAAGGAGACCTGAA 1860
DB 1801 CGGAAAGCAGAGATGTACGATTAACCTGTTGAAGCTTAAGATCAAGGAGACCTGAA 1860
QY 1861 GTGAGAGAAATGAGGTTGATGATCAAGTTGAGAGGATCTTGTCTCAGAGAGAA 1920
DB 1861 GTGAGAGAAATGAGGTTGATGATCAAGTTGAGAGGATCTTGTCTCAGAGAGAA 1920
QY 1921 CTCGAGAGAACTGAGAGAGGCTGTAAAGCCAGACAGAGGACACGAGCTGCTCAG 1980
DB 1921 CTCGAGAGAACTGAGAGAGGCTGTAAAGCCAGACAGAGGACACGAGCTGCTCAG 1980
QY 1981 AATATCCGCCAGGACAAAGAGACGAGCGAGAGGAGCTGAGAAAGCTTCAGAACCGAGAG 2040
DB 1981 AATATCCGCCAGGACAAAGAGACGAGCGAGAGGAGCTGAGAAAGCTTCAGAACCGAGAG 2040
QY 2041 GATTCCTTGAAGGATCAAGAAAGAGTGTGAGAGCTGAGAAAGCTGAGAAAGCTTCG 2100
DB 2041 GATTCCTTGAAGGATCAAGAAAGAGTGTGAGAGCTGAGAAAGCTGAGAAAGCTTCG 2100
QY 2101 GAGAAACAAGGTTAAAGAGCTAGAGACCATGAGAGGTTAGAGAAACAGACTGAAGATGAC 2160
DB 2101 GAGAAACAAGGTTAAAGAGCTAGAGACCATGAGAGGTTAGAGAAACAGACTGAAGATGAC 2160
QY 2161 ATCCAGACAAAAATCCCAACAGATCCAGAGATGCTGATTAATTTCTGAGCTGAAAGAG 2220
DB 2161 ATCCAGACAAAAATCCCAACAGATCCAGAGATGCTGATTAATTTCTGAGCTGAAAGAG 2220
QY 2221 AAAAATCGGGAGGCCCAAGTCTCAGCCACGACCTAGAAAGTGCACCTGMAACAGAAAGAG 2280
DB 2221 AAAAATCGGGAGGCCCAAGTCTCAGCCACGACCTAGAAAGTGCACCTGMAACAGAAAGAG 2280
QY 2281 CAGCACTATGAGAAAGATTAAAGTGTGAGCAATCAGATTAAGAAAGAGCTGAGTAC 2340
DB 2281 CAGCACTATGAGAAAGATTAAAGTGTGAGCAATCAGATTAAGAAAGAGCTGAGTAC 2340
QY 2341 AAGAGACACTGAGAAACATGATSCAGAGACAGAGAGAGAGGCCCATGAGAGAGGCAAA 2400
DB 2341 AAGAGACACTGAGAAACATGATSCAGAGACAGAGAGAGAGGCCCATGAGAGAGGCAAA 2400
QY 2401 ATTCTCAGCGAAAGAGGCGATGATCAATGCTATGATTCAAAGATCAAGTCCCTGAA 2460
DB 2401 ATTCTCAGCGAAAGAGGCGATGATCAATGCTATGATTCAAAGATCAAGTCCCTGAA 2460

QY 2461 CAGAGATTTGTGAACTGTCTGAAAGCCAAATTAACCTTCAGCAATATACAGCTTTTACC 2520
DB 2461 CAGAGATTTGTGAACTGTCTGAAAGCCAAATTAACCTTCAGCAATATACAGCTTTTACC 2520
QY 2521 CAAAGGAACTGAAGGCCCAAGAGAGATGATTTCTGAGCTCAGGCAACAGAAATTTTAC 2580
DB 2521 CAAAGGAACTGAAGGCCCAAGAGAGATGATTTCTGAGCTCAGGCAACAGAAATTTTAC 2580
QY 2581 CTGAGACACAGGCTGGAGAGTTGAGAGGCCAGAACCCGAAATCTGAGAGCAGCTGAG 2640
DB 2581 CTGAGACACAGGCTGGAGAGTTGAGAGGCCAGAACCCGAAATCTGAGAGCAGCTGAG 2640
QY 2641 AAGATCCACCAAGACCAAGTGAACAAGATTCGCTGCTGCAATCTGAGACAAAGTTG 2700
DB 2641 AAGATCCACCAAGACCAAGTGAACAAGATTCGCTGCTGCAATCTGAGACAAAGTTG 2700
QY 2701 CGGAGGTCAGTCTAGAGCAAGAGAGCAGAAACTGAGGCTCAAGGCCAGCTCAAGAG 2760
DB 2701 CGGAGGTCAGTCTAGAGCAAGAGAGCAGAAACTGAGGCTCAAGGCCAGCTCAAGAG 2760
QY 2761 CTACAGCTTCCCTGACAGAGCGGAGTCAAGTTGACAGCCCTGACAGGCTGACAGGCG 2820
DB 2761 CTACAGCTTCCCTGACAGAGCGGAGTCAAGTTGACAGCCCTGACAGGCTGACAGGCG 2820
QY 2821 GCCCTGAGAGCCAGCTTCCGACAGCGAGACAGAGCTGAGAGACCAAGCAAGAGCT 2880
DB 2821 GCCCTGAGAGCCAGCTTCCGACAGCGAGACAGAGCTGAGAGACCAAGCAAGAGCT 2880
QY 2881 GAAGAGAGATTCAGAGACCTCAAGGACATAGAGATTAATTCAGGCCCAATTTGAGCT 2940
DB 2881 GAAGAGAGATTCAGAGACCTCAAGGACATAGAGATTAATTCAGGCCCAATTTGAGCT 2940
QY 2941 CTTGCTAACAGCTGTACTGTATCAAGACTGAGAGAGCAGCTTAACAGCTGACCGAG 3000
DB 2941 CTTGCTAACAGCTGTACTGTATCAAGACTGAGAGAGCAGCTTAACAGCTGACCGAG 3000
QY 3001 GACACGCTGAACTCAACAAACCTTCTGCTCAACCAACTCGATGAGCTTCT 3060
DB 3001 GACACGCTGAACTCAACAAACCTTCTGCTCAACCAACTCGATGAGCTTCT 3060
QY 3061 GGGGCCCAAGCAGATTTGTAACACTGCGAAGTGAACCTATCCCGCGGAGATC 3120
DB 3061 GGGGCCCAAGCAGATTTGTAACACTGCGAAGTGAACCTATCCCGCGGAGATC 3120
QY 3121 ACCGAAAGAGATGACGCTTACCAAGCCAGAAAGCAGATGAGGCTTGAAGACCA 3180
DB 3121 ACCGAAAGAGATGACGCTTACCAAGCCAGAAAGCAGATGAGGCTTGAAGACCA 3180
QY 3181 TGACCAATGCTGAGAGAAACAGTTCATGATTTGAGAGCCCTAAACGATGAGCTGAGAA 3240
DB 3181 TGACCAATGCTGAGAGAAACAGTTCATGATTTGAGAGCCCTAAACGATGAGCTGAGAA 3240
QY 3241 AAAGAGCGGAGTGGAGGCGCTGAGAGAGCTCTGGGTGATGAGAAATCCCACTTGGAG 3300
DB 3241 AAAGAGCGGAGTGGAGGCGCTGAGAGAGCTCTGGGTGATGAGAAATCCCACTTGGAG 3300
QY 3301 TGTGAGGTTGAGAGCTGAGAGGATGCTGAGACACGAGAAACAGACAGAGGCGAGAGCC 3360
DB 3301 TGTGAGGTTGAGAGCTGAGAGGATGCTGAGACACGAGAAACAGACAGAGGCGAGAGCC 3360
QY 3361 GATCAGCGATCAACCGAGTCTCGCAGAGTGTGAGAGCTGACAGTGAAGAGACCAAGCT 3420
DB 3361 GATCAGCGATCAACCGAGTCTCGCAGAGTGTGAGAGCTGACAGTGAAGAGACCAAGCT 3420
QY 3421 GAGATTTCTGCTGTCAGAGGCTCTCAAGAGAGAGAGCTGAAGGCCGAGAGCTTCT 3480
DB 3421 GAGATTTCTGCTGTCAGAGGCTCTCAAGAGAGAGAGCTGAAGGCCGAGAGCTTCT 3480
QY 3481 GACAGCTCAATGACCTGAGAGAGAGACATGCTATGCTTGAATGAATGACCCGAAAGCTTA 3540
DB 3481 GACAGCTCAATGACCTGAGAGAGAGACATGCTATGCTTGAATGAATGACCCGAAAGCTTA 3540
QY 3541 CAGCAGAGCTGAGACTGAACGAGAGCTCAACAGAGGCTTCTGAGAGAGAGCCAAA 3600

Db	3541		CAGGAGAGCTGGAGACTGAACGAGAGCTCAAGAGGCTCTTGGAGAGACAGCCAAA	3600
Qy	3601		TTACAGCAGCAGATGAGACTTCAGAGAAAATATCAATTTTCGTCGTGACTCAAGACTGCA	3660
Db	3601		TTACAGCAGCAGATGAGACTTCAGAGAAAATATCAATTTTCGTCGTGACTCAAGACTGCA	3660
Qy	3661		GAGCTCTAGATGGGGCTGATCTACTAGAGCAGAAAAGATGACTTGGAGATACAGCTG	3720
Db	3661		GAGCTCTAGATGGGGCTGATCTACTAGAGCAGAAAAGATGACTTGGAGATACAGCTG	3720
Qy	3721		GAAGAACTTCAAGTTCATTTCTATTTCTCATGTAAAGGTGAAAATGAAAGGCACTATTCTCA	3780
Db	3721		GAAGAACTTCAAGTTCATTTCTATTTCTCATGTAAAGGTGAAAATGAAAGGCACTATTCTCA	3780
Qy	3781		CAAAACCAACTCATTTGATTTTCTGCAAGCCAAAATGAGCCAACTGTCTAAAGAAAAG	3840
Db	3781		CAAAACCAACTCATTTGATTTTCTGCAAGCCAAAATGAGCCAACTGTCTAAAGAAAAG	3840
Qy	3841		GTTCTCTGAGTACATGTAGCTGAGACTGCGCTCGAGAAAGGAAAGTCCGTGTGCA	3900
Db	3841		GTTCTCTGAGTACATGTAGCTGAGACTGCGCTCGAGAAAGGAAAGTCCGTGTGCA	3900
Qy	3901		GAGCTAGAGAGGCTTCAGAAAGCCGCAATCGAGCTCCGGTCGCGCGGAGAGAGT	3960
Db	3901		GAGCTAGAGAGGCTTCAGAAAGCCGCAATCGAGCTCCGGTCGCGCGGAGAGAGT	3960
Qy	3961		GCCCAACCGCAAGCAAGCAAGCAAGCCACATCCAGCCAGCAACGAGCAAGCAAGCAAG	4020
Db	3961		GCCCAACCGCAAGCAAGCAAGCAAGCCACATCCAGCCAGCAACGAGCAAGCAAGCAAG	4020
Qy	4021		ATGCGCATGTCGCCCATTCGTCGGTGGCCAGAGCAACAGCCAGTGCATGAGCTGTCTG	4080
Db	4021		ATGCGCATGTCGCCCATTCGTCGGTGGCCAGAGCAACAGCCAGTGCATGAGCTGTCTG	4080
Qy	4081		GCCCCCCTCATCCAGCCGAGAAAGAGTCTTCAACTCAGAGGAAATTTAGTCGGCGCTT	4140
Db	4081		GCCCCCCTCATCCAGCCGAGAAAGAGTCTTCAACTCAGAGGAAATTTAGTCGGCGCTT	4140
Qy	4141		AAGGAACGATGACACCAATATTTCTCAACCGATTCAACGATGAGCTGAACATGCAAGCC	4200
Db	4141		AAGGAACGATGACACCAATATTTCTCAACCGATTCAACGATGAGCTGAACATGCAAGCC	4200
Qy	4201		ACAAAGTGTGTCGT	4260
Db	4201		ACAAAGTGTGTCGT	4260
Qy	4261		GATGTCAAGT	4320
Db	4261		GATGTCAAGT	4320
Qy	4321		CCTGCTGAATATGCAACACATTTCAACGAGGCTTGTGCGTGAACAAAATGAATCTCCCA	4380
Db	4321		CCTGCTGAATATGCAACACATTTCAACGAGGCTTGTGCGTGAACAAAATGAATCTCCCA	4380
Qy	4381		GGTCTCCAGACCAAGAGCCCAAGCAGACTTGTGCACTTGAAGGTGAGTGAAGTGTGCC	4440
Db	4381		GGTCTCCAGACCAAGAGCCCAAGCAGACTTGTGCACTTGAAGGTGAGTGAAGTGTGCC	4440
Qy	4441		AGGAATTAACAAAGAGCAGCAAGGCTGGGACAGGAATGATCTCTGAGAGGATCA	4500
Db	4441		AGGAATTAACAAAGAGCAGCAAGGCTGGGACAGGAATGATCTCTGAGAGGATCA	4500
Qy	4501		AAAGTCTCTATTTATGACATGTAAGCAGAGAGCTGACAGAGCCGCTGAGAAATTT	4560
Db	4501		AAAGTCTCTATTTATGACATGTAAGCAGAGAGCTGACAGAGCCGCTGAGAAATTT	4560
Qy	4561		GAGCTGTGCTTCCGAGCGGGATGTATCTATTTCAATGTGTGCTGTGTGTGTGTGTGT	4620
Db	4561		GAGCTGTGCTTCCGAGCGGGATGTATCTATTTCAATGTGTGCTGTGTGTGTGTGTGT	4620
Qy	4621		GCAATTAACGCAAGAGAGTGTCCATATACATGAGAGTGAATTTCAACCCGACAC	4680
Db	4621		GCAATTAACGCAAGAGAGTGTCCATATACATGAGAGTGAATTTCAACCCGACAC	4680
Db	4681		ACCTGTGAGCCGAGAGAACCTCTACTTGTCTAGCTCCAGACTTCCCTGACAAACAGCGC	4740
Qy	4741		TGGGTCAACCCCTTGAATCAATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	4800
Db	4741		TGGGTCAACCCCTTGAATCAATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	4800
Qy	4801		GCTGATCTTAACTGCTTGAACCTCTCTGTAACCTGTAAGGTGATGACCTGTAGAC	4860
Db	4801		GCTGATCTTAACTGCTTGAACCTCTCTGTAACCTGTAAGGTGATGACCTGTAGAC	4860
Qy	4861		ATGAATGCAAGCTGTGCTTCAATGACAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGT	4920
Db	4861		ATGAATGCAAGCTGTGCTTCAATGACAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGT	4920
Qy	4921		TACGCTGATGTCTTGAAGAACTCCCTTAACCATGTGCCAGAGAAATGGAGCATCTTC	4980
Db	4921		TACGCTGATGTCTTGAAGAACTCCCTTAACCATGTGCCAGAGAAATGGAGCATCTTC	4980
Qy	4981		CAAAATTAATTAATTAACAGGACCTGAGAGGCTTCAATGATAGCAGAGAGAGCGGCA	5040
Db	4981		CAAAATTAATTAATTAACAGGACCTGAGAGGCTTCAATGATAGCAGAGAGAGCGGCA	5040
Qy	5041		CTGTGTCTGTGAGCGTGAAGAAAGTGAACAGTCCCTGGCCAGTCCCACTGCTGCC	5100
Db	5041		CTGTGTCTGTGAGCGTGAAGAAAGTGAACAGTCCCTGGCCAGTCCCACTGCTGCC	5100
Qy	5101		CAGCCGCAATCTCACCCCAATTTTGAAGCTGTCAAGAGGCTGCCATTTGTTGGGGCA	5160
Db	5101		CAGCCGCAATCTCACCCCAATTTTGAAGCTGTCAAGAGGCTGCCATTTGTTGGGGCA	5160
Qy	5161		GGCAAGATTGAAGAGGAGCTGTGACATCTGTGACCAATGCCAGAGAAATCTGATTTCT	5220
Db	5161		GGCAAGATTGAAGAGGAGCTGTGACATCTGTGACCAATGCCAGAGAAATCTGATTTCT	5220
Qy	5221		CGCTAACAGAAACCTCAGCAAAATCTGATCCGAGAAAGATAGAGACTTCAGAGCC	5280
Db	5221		CGCTAACAGAAACCTCAGCAAAATCTGATCCGAGAAAGATAGAGACTTCAGAGCC	5280
Qy	5281		TGAGCTGTATCTCACTTCAACCAATTAAGTATCTCTCAATTTGAACCAATTAATTTA	5340
Db	5281		TGAGCTGTATCTCACTTCAACCAATTAAGTATCTCTCAATTTGAACCAATTAATTTA	5340
Qy	5341		ATGCAATGAAGCAGTACAGCTGAGAGAAATTCGAGTAAGAAATGAACCAATTTCTGGA	5400
Db	5341		ATGCAATGAAGCAGTACAGCTGAGAGAAATTCGAGTAAGAAATGAACCAATTTCTGGA	5400
Qy	5401		CCTGTGTGTGTGCGGCTCTTCAACAGCTTCCCTGTCTCAATGTGTGAGGTGAACAGC	5460
Db	5401		CCTGTGTGTGTGCGGCTCTTCAACAGCTTCCCTGTCTCAATGTGTGAGGTGAACAGC	5460
Qy	5461		GCAAGGACAGGAGAGAGTACTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	5520
Db	5461		GCAAGGACAGGAGAGAGTACTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	5520
Qy	5521		TACGGAAGACGTAGGCGGACAGACGATCTCAAGTGAAGTGTGTGTGTGTGTGTGTGT	5580
Db	5521		TACGGAAGACGTAGGCGGACAGACGATCTCAAGTGAAGTGTGTGTGTGTGTGTGTGT	5580
Qy	5581		TACAGAGAACCTTACTGT	5640
Db	5581		TACAGAGAACCTTACTGT	5640
Qy	5641		GCAAGCTCTGAGAGAGACCTGTGCCAGAGTACTTGAACATCCGAAACCGGCTTAC	5700
Db	5641		GCAAGCTCTGAGAGAGACCTGTGCCAGAGTACTTGAACATCCGAAACCGGCTTAC	5700
Qy	5701		CTGGGCTCTGCAATTTCTCAGAGCCGATTTACTTGGGCTCTCATACAGATTAATTA	5760
Db	5701		CTGGGCTCTGCAATTTCTCAGAGCCGATTTACTTGGGCTCTCATACAGATTAATTA	5760

Qy 5761 AGGGTCATTTGTCGCAAGGAAACCTGTGAAGAGTCCGGCACTGAACCAACCGGGGC 5820
Db 5761 AGGGTCATTTGTCGCAAGGAAACCTGTGAAGAGTCCGGCACTGAACCAACCGGGGC 5820
Qy 5821 CCGTCCACCTCCCGCAG 5837
Db 5821 CCGTCCACCTCCCGCAG 5837

RESULT 11
US-10-017-216-3
Sequence 3, Application US/10017216
Publication No. US20020160483A1
GENERAL INFORMATION:
APPLICANT: KABELLER-LIBERMANN, Rosana
TITLE OF INVENTION: 13245, A No. US20020160483A1el Human Myotonic Dystrophy Type Prot
FILE REFERENCE: 10147-5701
CURRENT APPLICATION NUMBER: US/10/017,216
CURRENT FILING DATE: 2001-10-23
PRIOR APPLICATION NUMBER: US 60/242,429
PRIOR FILING DATE: 2000-10-23
NUMBER OF SEQ ID NOS: 7
SOFTWARE: Patent In Ver. 2.1
SEQ ID NO 3
LENGTH: 6159
TYPE: DNA
ORGANISM: Homo sapiens
US-10-017-216-3

Query Match 91.8%; Score 5661.4; DB 13; Length 6159;
Best Local Similarity 95.8%; Pred. No. 0;
Matches 5944; Conservative 0; Mismatches 16; Indels 243; Gaps 3;

Qy 1 ATGTTGAAGTTCAAAATATGAGCGCGGAATCCTTTGATGCTGTGCTGTAACCAATT 60
Db 1 ATGTTGAAGTTCAAAATATGAGCGCGGAATCCTTTGATGCTGTGCTGTAACCAATT 60
Qy 61 GCCACCGCGGCTCCAGGCTGAATCTGTTCCAGGAGAAACCAACCTTTATGACTCAA 120
Db 61 GCCACCGCGGCTCCAGGCTGAATCTGTTCCAGGAGAAACCAACCTTTATGACTCAA 120
Qy 121 CACGAGATGTCCTCTTTCGCCGAGAGATTTAGATGCCCTTTGTTCTTTGAA 180
Db 121 CACGAGATGTCCTCTTTCGCCGAGAGATTTAGATGCCCTTTGTTCTTTGAA 180
Qy 181 GAATGAGTCAAGCTGCTGTGATGAAGATTAAAGCACTGAGCAACTTTGTCGGAGTAT 240
Db 181 GAATGAGTCAAGCTGCTGTGATGAAGATTAAAGCACTGAGCAACTTTGTCGGAGTAT 240
Qy 241 TCCGACACCATAGTGAATTAAGAGAGTCCAGCTTCGGAGAGACTTGAAGTCA 300
Db 241 TCCGACACCATAGTGAATTAAGAGAGTCCAGCTTCGGAGAGACTTGAAGTCA 300
Qy 301 AGTCTTGAAGTTGTGCTCACTTGTGAGAGTGAAGTGAAGAGAAACCAACCGGG 360
Db 301 AGTCTTGAAGTTGTGCTCACTTGTGAGAGTGAAGTGAAGAGAAACCAACCGGG 360
Qy 361 GACATCTATGCTATGAAGATGATGAAGAGGCTTTATTTGGCCAGAGAGGTTTCA 420
Db 361 GACATCTATGCTATGAAGATGATGAAGAGGCTTTATTTGGCCAGAGAGGTTTCA 420
Qy 421 TTTTGTGAGAGAGCGGAACTATTTATCTGAAGCAACAGCCGTGATCCCAATTA 480
Db 421 TTTTGTGAGAGAGCGGAACTATTTATCTGAAGCAACAGCCGTGATCCCAATTA 480
Qy 481 CAGTATGCTTTGAGAGCAAAATCACTTTATCTGATGAGAGATATCAAGCTGAGGG 540
Db 481 CAGTATGCTTTGAGAGCAAAATCACTTTATCTGATGAGAGATATCAAGCTGAGGG 540
Qy 541 GACTTGTGCTCACTTTGAATGATATGAGAGCAAGTTAGTGAACCTGATACAGTTT 600
Db 541 GACTTGTGCTCACTTTGAATGATATGAGAGCAAGTTAGTGAACCTGATACAGTTT 600

Db 541 GACTTGTGCTCACTTTGAATGATATGAGAGCAAGTTAGTGAACCTGATACAGTTT 600
Qy 601 TACCTAGCTGAGTGAATTTTGGCTGTTCACAGGCTTCATCTGATGAGATACGTGATCA 660
Db 601 TACCTAGCTGAGTGAATTTTGGCTGTTCACAGGCTTCATCTGATGAGATACGTGATCA 660
Qy 661 GACATCAAGCTTGAAGCAATTTCTGTTGACCCGACAGAGACATTCAGCTGAGGATTTT 720
Db 661 GACATCAAGCTTGAAGCAATTTCTGTTGACCCGACAGAGACATTCAGCTGAGGATTTT 720
Qy 721 GATCTGCGCGGAAATGAATTCAAACAGATGATGATCCAAATCCCGATTTGGAGC 780
Db 721 GATCTGCGCGGAAATGAATTCAAACAGATGATGATCCAAATCCCGATTTGGAGC 780
Qy 781 CCAGATTACATGCTCCTGTAAGTCTGATGATGAACGGGATGAGAAAGGACCTTAC 840
Db 781 CCAGATTACATGCTCCTGTAAGTCTGATGATGAACGGGATGAGAAAGGACCTTAC 840
Qy 841 GGCCTGAGCTGTACTGTGTGTGATGAGGCGGTGATTCCTATGATGATTTATGGAGA 900
Db 841 GGCCTGAGCTGTACTGTGTGTGATGAGGCGGTGATTCCTATGATGATTTATGGAGA 900
Qy 901 TCCCGCTTGCAGAGGGAACCTGTCAGAACCTTCAATTAATTAAGAAATTTCCACGG 960
Db 901 TCCCGCTTGCAGAGGGAACCTGTCAGAACCTTCAATTAATTAAGAAATTTCCACGG 960
Qy 961 TTTTGAATTTTCCAGATGACCCCAAGTGAAGTGAAGTCTTTGATCTGATTTCAAGC 1020
Db 961 TTTTGAATTTTCCAGATGACCCCAAGTGAAGTGAAGTCTTTGATCTGATTTCAAGC 1020
Qy 1021 TTGTTGTGCGGCAAGAAAGAGACTGAAGTTGAAGTCTTTGCTGCAATCTTTCTTC 1080
Db 1021 TTGTTGTGCGGCAAGAAAGAGACTGAAGTTGAAGTCTTTGCTGCAATCTTTCTTC 1080
Qy 1081 TCTTAAATTAATCTGGAACCAATTCGTAATCTCTCCCGCTTCTGTCACCCCTGAG 1140
Db 1081 TCTTAAATTAATCTGGAACCAATTCGTAATCTCTCCCGCTTCTGTCACCCCTGAG 1140
Qy 1141 TCTGACGATGACACTTCAATTTTGAATGAACAGAGAAATTCGTTGCTTCTCT 1200
Db 1141 TCTGACGATGACACTTCAATTTTGAATGAACAGAGAAATTCGTTGCTTCTCT 1200
Qy 1201 CCGTGCAGCTGAGACCCCTCAGGCTTCTCGGATGAAGACTGCCGTTTGTGGGTTTCG 1260
Db 1201 CCGTGCAGCTGAGACCCCTCAGGCTTCTCGGATGAAGACTGCCGTTTGTGGGTTTCG 1260
Qy 1261 TACAGCAAGCACTGGGATTTCTTGTGATCTGAGTCTGTGCTGAGGCTTGGACTCC 1320
Db 1261 TACAGCAAGCACTGGGATTTCTTGTGATCTGAGTCTGTGCTGAGGCTTGGACTCC 1320
Qy 1321 CTTGCCAAGTACTAGCTCATGAAAGAACTTCTCATCAAAAGCAAGAGCTTCAAGAC 1380
Db 1321 CTTGCCAAGTACTAGCTCATGAAAGAACTTCTCATCAAAAGCAAGAGCTTCAAGAC 1380
Qy 1381 TCTCAGGACAAGTGTCACAAGATGAGACAGAAATGAACCGGTTTACATCGAGAGTCTCA 1440
Db 1381 TCTCAGGACAAGTGTCACAAGATGAGACAGAAATGAACCGGTTTACATCGAGAGTCTCA 1440
Qy 1441 GAGGTGAGGCTGTGCTTATGATCAGAGAGAGTGAAGGCTTGTGAGACTCAAGAGA 1500
Db 1441 GAGGTGAGGCTGTGCTTATGATCAGAGAGAGTGAAGGCTTGTGAGACTCAAGAGA 1500
Qy 1501 TCCCTCTGAGAGAGGACTTGTGCTTACATCATCAAGATGAGAGTCACTTAAAGAGT 1560
Db 1501 TCCCTCTGAGAGAGGACTTGTGCTTACATCATCAAGATGAGAGTCACTTAAAGAGT 1560
Qy 1561 TTGAGAGCAAGCAAGATGAGAGTGTCCAGAGAGATCAAAAGCACTGACGTTCTCAT 1620
Db 1561 TTGAGAGCAAGCAAGATGAGAGTGTCCAGAGAGATCAAAAGCACTGACGTTCTCAT 1620
Qy 1621 GATATCAGAGAGCAAGCCGAGAGCTTCAAGAAATCAAGAGCAAGGATACAGGCTCAA 1680
Db 1621 GATATCAGAGAGCAAGCCGAGAGCTTCAAGAAATCAAGAGCAAGGATACAGGCTCAA 1680

1681 GTGGAAGAAATGAGTTGATGATGAAATGAGTTGGAAGGATCTTGTCTCAGCAAGAA 1740
1681 GTGGAAGAAATGAGTTGATGATGAAATGAGTTGGAAGGATCTTGTCTCAGCAAGAA 1740
1741 CGGAGTGTCTCTAGAACTGAGCTGAGAGAGTCTCGGCTTGGCTGAGGAATTTACG 1800
1741 CGGAGTGTCTCTAGAACTGAGCTGAGAGAGTCTCGGCTTGGCTGAGGAATTTACG 1800
1801 CGGAAGCGACAGATGTGACGATTAACCTGTTGAAGGCTTAAGATCAAGGAGAGCTGAA 1860
1801 CGGAAGCGACAGATGTGACGATTAACCTGTTGAAGGCTTAAGATCAAGGAGAGCTGAA 1860
1861 GTGGGAAGAAATGAGAACTGGAAGAAATCAATCTGAGCAGAGCTCAAAATTCAGAG 1920
1861 GTGGGAAGAAATGAGAACTGGAAGAAATCAATCTGAGCAGAGCTCAAAATTCAGAG 1920
1921 CTCGAAGAGAACTGAGAAAGGCTGTAAAGCCAGCAGCAGAGCCAGAGCTGCTCAG 1980
1921 CTCGAAGAGAACTGAGAAAGGCTGTAAAGCCAGCAGCAGAGCCAGAGCTGCTCAG 1980
1981 AATATTCGCGACGCAAGGAGCCGAGAGGAGCTGAGAGAGCTGAGAGAGCCGAGAG 2040
1981 AATATTCGCGACGCAAGGAGCCGAGAGGAGCTGAGAGAGCTGAGAGAGCCGAGAG 2040
1945 -----GCAAGGAGCAGCCGAGAGGAGCTGAGAGAGCTGAGAGAGCCGAGAG 1992
2041 GATTCTTCTGAAGGCAATCAAGAAAGCTGTGAGAGCTGAGAGAGCCGCTATTCTTG 2100
2041 GATTCTTCTGAAGGCAATCAAGAAAGCTGTGAGAGCTGAGAGAGCCGCTATTCTTG 2100
1993 GATTCTTCTGAAGGCAATCAAGAAAGCTGTGAGAGCTGAGAGAGCCGCTATTCTTG 2052
2101 GAGAACAGGTAAAGAGACTAGAGACCATGAGAGCTGAGAGAAACAGACTGAAGAGATGAC 2160
2053 GAGAACAGGTAAAGAGACTAGAGACCATGAGAGCTGAGAGAAACAGACTGAAGAGATGAC 2112
2161 ATTCAGACAAATATCCCAACAGATCCAGCAGATGGCTGATTAATTTCTGAGCTCGAAGAG 2220
2113 ATTCAGACAAATATCCCAACAGATCCAGCAGATGGCTGATTAATTTCTGAGCTCGAAGAG 2172
2221 AAACATTCGGAGAGGCCCAAGTCTAGAGCCAGACCTTAGAAGTGCACCTGAAACAGAAAGAG 2280
2173 AAACATTCGGAGAGGCCCAAGTCTAGAGCCAGACCTTAGAAGTGCACCTGAAACAGAAAGAG 2232
2281 CAGCACTATGAGGAAAGATTAAGTGAACATTCAGATTAAGAAAGCCTGGCTGAG 2340
2233 CAGCACTATGAGGAAAGATTAAGTGAACATTCAGATTAAGAAAGCCTGGCTGAG 2292
2341 AAGGAGACACTGAGAGAACTGATGAGAGACACGAGAGAGAGGCCCATGAGAGAGCCAAA 2400
2293 AAGGAGACACTGAGAGAACTGATGAGAGACACGAGAGAGAGGCCCATGAGAGAGCCAAA 2352
2401 AATTCAGAGGAAACAGAAAGCGATGATCAATGCTATGGAATTCAGAGTCAAGATCCCTGAA 2460
2353 AATTCAGAGGAAACAGAAAGCGATGATCAATGCTATGGAATTCAGAGTCAAGATCCCTGAA 2412
2461 CAGAGGATTTGAGAACTGTCTGAGAGCAATTAACCTTGAGCAAAATGAGCTCTTTTAC 2520
2413 CAGAGGATTTGAGAACTGTCTGAGAGCAATTAACCTTGAGCAAAATGAGCTCTTTTAC 2472
2521 CAAAGGAACATGAGAGCCCAAGAAAGATGATTTCTGAACTCAGAGCAAGAAATTTTAC 2580
2473 CAAAGGAACATGAGAGCCCAAGAAAGATGATTTCTGAACTCAGAGCAAGAAATTTTAC 2532
2561 CTGGAGACACAGGCTGGAGAGTTGAGAGCCCAAGAACTGAGAGAGAGAGCTGAG 2640
2533 CTGGAGACACAGGCTGGAGAGTTGAGAGCCCAAGAACTGAGAGAGAGAGCTGAG 2592
2641 AAGATCAGACACAGAGCCCAAGAAAGATGATTTCTGAACTCAGAGCAAGAAATTTTAC 2700
2533 AAGATCAGACACAGAGCCCAAGAAAGATGATTTCTGAACTCAGAGCAAGAAATTTTAC 2652
2701 CGGAGAGTCAAGTCTAGAGACAGAGAGAGAACTGAGAGCTCAAGCCAGAGCTCAAGAG 2760
2653 CGGAGAGTCAAGTCTAGAGACAGAGAGAGAACTGAGAGCTCAAGCCAGAGCTCAAGAG 2712
2761 CTACAGCTTCTCCTGAGAGAGCCGAGTCAAGATTGACAGCCCTGAGAGCTGACCGGCG 2820
2713 CTACAGCTTCTCCTGAGAGAGCCGAGTCAAGATTGACAGCCCTGAGAGCTGACCGGCG 2772
2821 GCCCTGAGAGCCAGCTTCCGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGCT 2880
2773 GCCCTGAGAGCCAGCTTCCGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGCT 2832
2881 GAAAGAGAGATTCAGAGCACTCAGAGCAATAGAGATGAATTCAGAGAGAGAGAGAGCT 2940
2833 GAAAGAGAGATTCAGAGCACTCAGAGCAATAGAGATGAATTCAGAGAGAGAGAGAGCT 2892
2941 CTTCGTAACAGCTGTATCTGTATCAAGACCTGAGAGAGAGAGAGAGAGAGAGAGCT 3000
2893 CTTCGTAACAGCTGTATCTGTATCAAGACCTGAGAGAGAGAGAGAGAGAGAGAGCT 2952
3001 GACACGCTGAATCAACCAACCAAACTTACTTGTCCAAACATCTGATGAGAGCTTCT 3060
2953 GACACGCTGAATCAACCAACCAAACTTACTTGTCCAAACATCTGATGAGAGCTTCT 3012
3061 GGGCCCAACGAGAGATTGTAACATGCGAAGGAGAGTGAACATCTCCGCGGAGAGATC 3120
3013 GGGCCCAACGAGAGATTGTAACATGCGAAGGAGAGTGAACATCTCCGCGGAGAGATC 3072
3121 ACGGAAAGAGAGATGAGCTTACCAACGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3180
3073 ACGGAAAGAGAGATGAGCTTACCAACGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3132
3181 TGCACATCTGAG 3240
3133 TGCACATCTGAG 3192
3241 AAAGAGCCGAG 3300
3193 AAAGAGCCGAG 3252
3301 TGTGGGTTTGAAGAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3360
3253 TGTGGGTTTGAAGAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3312
3361 GATCAGCGGATCAACGAGTCTGCGCAGAGTGTGAGAGCTGAGAGAGAGAGAGAGAG 3420
3313 GATCAGCGGATCAACGAGTCTGCGCAGAGTGTGAGAGCTGAGAGAGAGAGAGAGAG 3372
3421 GAGATTCTGCTCTGAG 3480
3373 GAGATTCTGCTCTGAG 3432
3481 GACAAAGCTCAATGAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3540
3433 GACAAAGCTCAATGAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3492
3541 CAGCAGAAAGCTGAG 3600
3493 CAGCAGAAAGCTGAG 3552
3601 TTACAG 3660
3553 TTACAG 3612
3661 GAAAGCTTGAATCGGAGCTGATCTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3720
3613 GAAAGCTTGAATCGGAGCTGATCTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3672
3721 GAAAGCAATTCAGAGTTCTTATCTCATGAAAGAGAGAGAGAGAGAGAGAGAGAGAG 3780
3673 GAAAGCAATTCAGAGTTCTTATCTCATGAAAGAGAGAGAGAGAGAGAGAGAGAGAG 3732
3781 CAAACCAAACTCATGATTTTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3838
3733 CAAACCAAACTCATGATTTTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3792
3839 -----AGATTCTCTGCAATAC 3855

Db 3793 GGTTTATTAGTCGACGGAAAGGACCCTTGCTTACCACACAGGTTCTCTGCACTAC 3852
Qy 3856 AATGAGCTGAAGCTGGCCCTGAGAAAGAAAGCTGCTGTGCAAGCTAGAGAAACC 3915
Db 3853 AATGAGCTGAAGCTGGCCCTGAGAAAGAAAGCTGCTGTGCAAGCTAGAGAAACC 3912
Qy 3916 CTTGAGAAAGCCGCGATCGAGCTCCGCTCCGCGAGAGAGTGTCCCAACCGAAAGCA 3975
Db 3913 CTTGAGAAAGCCGCGATCGAGCTCCGCTCCGCGAGAGAGTGTCCCAACCGAAAGCA 3972
Qy 3976 ACGGACACCCACACCCATCCAGCGCAGCACCAGGAGAGAGATGCGCATGTCCGCC 4035
Db 3973 ACGGACACCCACACCCATCCAGCGCAGCACCAGGAGAGAGATGCGCATGTCCGCC 4032
Qy 4036 ATGCGCGCTGCGCAGAGACACAGCCCGATGCGCATGTGCTGCGCCCGCATCCAGC 4095
Db 4033 ATGCGCGCTGCGCAGAGACACAGCCCGATGCGCATGTGCTGCGCCCGCATCCAGC 4092
Qy 4096 CGCAGAAAGAGTCTTCACTCCAGAGAAATTTAGTGGCGCTTTAAGAAACGCATGCAC 4155
Db 4093 CGCAGAAAGAGTCTTCACTCCAGAGAAATTTAGTGGCGCTTTAAGAAACGCATGCAC 4152
Qy 4156 CACAAATTTCTCAACCGATTTCAAGTAGAGCTGAACATGAGAGCAGAAAGTGTCTGTG 4215
Db 4153 CACAAATTTCTCAACCGATTTCAAGTAGAGCTGAACATGAGAGCAGAAAGTGTGTG 4212
Qy 4216 TGTGTGATACCGTGCACTTTGAGCGCAGGACATCCAAATGTCTCGAATGTCAAGTATG 4275
Db 4213 TGTGTGATACCGTGCACTTTGAGCGCAGGACATCCAAATGTCTCGAATGTCAAGTATG 4272
Qy 4276 TGTCAACCCCAAGTCTCAGGTCTTGCAGCCACCTTGCGGCTTGTCTGTAATATGCC 4335
Db 4273 TGTCAACCCCAAGTCTCAGGTCTTGCAGCCACCTTGCGGCTTGTCTGTAATATGCC 4332
Qy 4336 ACACACTTTCACCCAGGCTTCTGCGGTGACAAATGAACTCTCCAGGTTCTCCAGACCAAG 4395
Db 4333 ACACACTTTCACCCAGGCTTCTGCGGTGACAAATGAACTCTCCAGGTTCTCCAGACCAAG 4392
Qy 4396 GAGCCCGACAGACGCTTGACCTGGAAGGTGATGAAAGGTGCCAGAAATACAAACGA 4455
Db 4393 GAGCCCGACAGACGCTTGACCTGGAAGGTGATGAAAGGTGCCAGAAATACAAACGA 4452
Qy 4456 GGAACAGCAAGCTGGGACAGGAAGTACATTTCTCTGAGGGATCAAAAGTCTCATTTAT 4515
Db 4453 GGAACAGCAAGCTGGGACAGGAAGTACATTTCTCTGAGGGATCAAAAGTCTCATTTAT 4512
Qy 4516 GACAAATGAGCAGAGAAAGCTGACAGAGCCGCTGGAAGATTTGAGCTGTGCCCTTCC 4575
Db 4513 GACAAATGAGCAGAGAAAGCTGACAGAGCCGCTGGAAGATTTGAGCTGTGCCCTTCC 4572
Qy 4576 GAGCGGATGATATTAATTCATGAGTCCGCTGAGTCTTCCGAATCCGAATACAGCCAAA 4635
Db 4573 GAGCGGATGATATTAATTCATGAGTCCGCTGAGTCTTCCGAATCCGAATACAGCCAAA 4632
Qy 4636 GCAGATGTCCATCATACTGAAGATGAAATCTCACCCGACACACCTGTGCGCCGGG 4695
Db 4633 GCA----- 4635
Qy 4696 AGAACCTCTAATTGAGTCTCCAGCTTCCCTGACAAACAGCGCTGGGTCAACCGCTTAA 4755
Db 4636 ----- 4635
Qy 4756 GAATCATTTGTGCGAGGTGGGAGAGTTTCTAGGAAAAAGCAGAAAGCTGATCTAACTG 4815
Db 4636 -----GAAAAAGCAAGAGCTGATGCTTAACTG 4662
Qy 4816 CTTGAAAATCTCTGCTGAAACTGAAAGGTGATGACCTCTAGACATGAACCTGACGCTG 4875
Db 4663 CTTGAAAATCTCTGCTGAAACTGAAAGGTGATGACCTCTAGACATGAACCTGACGCTG 4722
Qy 4876 CCCTTCAGTGAACAGGTGTGTGTGTGGGCAACGAGAAAGGCTCTACGCTTGAATGTC 4935
|||||

Db 4723 CCCTTCAGTGAACAGGTGTGTGTGTGGGCAACGAGAAAGGCTCTACGCTTGAATGTC 4782
Qy 4936 TTGAAAAATCTCTTAACCCATGTCCTCCAGAAATTTGAGCACTCTTCCAAATTTATATATC 4995
Db 4783 TTGAAAAATCTCTTAACCCATGTCCTCCAGAAATTTGAGCACTCTTCCAAATTTATATATC 4842
Qy 4996 AAGGACCTGGAGAAAGCTACTATGATAGCAGAGAAAGAGCGGCACTGTGTCTTGTGAC 5055
Db 4843 AAGGACCTGGAGAAAGCTACTATGATAGCAGAGAAAGAGCGGCACTGTGTCTTGTGAC 4992
Qy 5056 GTGAAAGAAAGTGAACACAGTCCCTGAGCCAGTCCACCTGCTGCCAGCCGACATCTCA 5115
Db 4903 GTGAAAGAAAGTGAACACAGTCCCTGAGCCAGTCCACCTGCTGCCAGCCGACATCTCA 4962
Qy 5116 CCCAACATTTTGAAGCTGTCAAGAGCTGCCACTGTTGTGGGCAAGCAAGATTTGAAGAC 5175
Db 4963 CCCAACATTTTGAAGCTGTCAAGAGCTGCCACTGTTGTGGGCAAGCAAGATTTGAAGAC 5022
Qy 5176 GGGCTCTGCAATCTGTGAGGCAATGCCAGAAAGTGTCTTCTCGGCTCAACGAAAC 5235
Db 5023 GGGCTCTGCAATCTGTGAGGCAATGCCAGAAAGTGTCTTCTCGGCTCAACGAAAC 5082
Qy 5236 CTGACCAATACTGATCCGGAAGAGATAGACCTCAGAGCCCTGACCTGTATCCAC 5295
Db 5083 CTGACCAATACTGATCCGGAAGAGATAGACCTCAGAGCCCTGACCTGTATCCAC 5142
Qy 5296 TTGACCAATTTACAGTATCTCTATTGGAACCAATTAATTTCAACAAATTCAGATGACAG 5355
Db 5143 TTGACCAATTTACAGTATCTCTATTGGAACCAATTAATTTCAACAAATTCAGATGACAG 5202
Qy 5356 TACACGCTCGAGAAATTTCTGATTAAGAAATGACATTTCTTGGACCTGTGTGTGGCC 5415
Db 5203 TACACGCTCGAGAAATTTCTGATTAAGAAATGACATTTCTTGGACCTGTGTGTGGCC 5282
Qy 5416 GCCCTTTCACAAAGCTTCCCTGTCTCAATGATGAGTGAACAGCGAGGACAGAGAG 5475
Db 5263 GCCCTTTCACAAAGCTTCCCTGTCTCAATGATGAGTGAACAGCGAGGACAGAGAG 5322
Qy 5476 GAGTACTTGTGTGTGTTCACAGAAATTTGAGTGTCTGTGATTTCTTACGAAAGAGTATG 5535
Db 5323 GAGTACTTGTGTGTGTTCACAGAAATTTGAGTGTCTGTGATTTCTTACGAAAGAGTATG 5382
Qy 5536 CGCAGACGATCTCAAGTGAAGTCCGCTTACCTTTGGCTTTGTGCTTACAGAAACCTCAT 5595
Db 5383 CGCAGACGATCTCAAGTGAAGTCCGCTTACCTTTGGCTTTGTGCTTACAGAAACCTCAT 5442
Qy 5596 CTGTTGTGACCACTTCACTCACTGAGAGTATTTGAGTTCAGGAGCGCTCTCAACA 5655
Db 5443 CTGTTGTGACCACTTCACTCACTGAGAGTATTTGAGTTCAGGAGCGCTCTCAACA 5582
Qy 5656 GGGACCCCTGCGGAGGCTACTGAGATCCCGAACCCGCGCTTACTTGGGCTTGCATT 5715
Db 5503 GGGACCCCTGCGGAGGCTACTGAGATCCCGAACCCGCGCTTACTTGGGCTTGCATT 5552
Qy 5716 TCTCTAGAGAGATTTAATTGAGTCTCTCAATACAGAGTAAATTAAGGTCATTTGCTGC 5775
Db 5563 TCTCTAGAGAGATTTAATTGAGTCTCTCAATACAGAGTAAATTAAGGTCATTTGCTGC 5622
Qy 5776 AAGGAAAACCTCTGAGAGAGTCCGCGACITGAACACACCGGGGCGCTCCACTCCGC 5835
Db 5623 AAGGAAAACCTCTGAGAGAGTCCGCGACITGAACACACCGGGGCGCTCCACTCCGC 5682
Qy 5836 AGCAGCCCAACAGAGGCGCCACACGTAACACAGACATCACCAAGCGGTGCGC 5895
Db 5683 AGCAGCCCAACAGAGGCGCCACACGTAACACAGAGCAATCACCAAGCGGTGCGC 5742
Qy 5896 TCCAGCCAGGCGCGCGCGAGAGCCCAAGCCACCTCGGAGAGCCAGACACCCACCGC 5955
Db 5743 TCCAGCCCAAGCGCGCGCGAGAGCCCAAGCCACCTCGGAGAGCCCAAGCCACCGC 5802
Qy 5956 TACCGCAAGGGGCGAGACCGAGCTGCGGAGGAGCAAGTCTCTGCGCGCCCTCTGAGAGGA 6015
Db 5803 TACCGCAAGGGGCGAGACCGAGCTGCGGAGGAGCAAGTCTCTGCGCGCCCTCTGAGAGGA 5862

QY 6016 GAGAAATCCCCCGGCGGATCTAGCAACGGGAGAGAGCGGCTCCCGCGAGGCTGTTT 6075
DB 5863 GAGAAATCCCCCGGCGGATCTAGCAACGGGAGAGAGCGGCTCCCGCGAGGCTGTTT 5922
QY 6076 GAAGACAGACAGAGGCGGCTGCTGCGGAGACCGGAGGAGACCGGCTGCTGCGGAGG 6135
DB 5923 GAAGACAGACAGAGGCGGCTGCTGCGGAGACCGGAGGAGACCGGCTGCTGCGGAGG 5982
QY 6136 AACAAAGTCTGGAGCAAGCTTC 6158
DB 5983 AACAAAGGAGAGGCGAGAGTGC 6005

RESULT 12

US-10-325-430-11
Sequence 11, Application US/10325430
Publication No. US20030153525A1
GENERAL INFORMATION:
APPLICANT: Millennium Pharmaceuticals, Inc
APPLICANT: Siles-Santiago, Immaculada
APPLICANT: Rosenfeld, Julie Beth
TITLE OF INVENTION: METHODS AND COMPOSITIONS IN TREATING
TITLE OF INVENTION: PAIN AND PAINFUL DISORDERS USING 1465, 1587, 2146, 2207,
FILE REFERENCE: MP101-294P1RNM
CURRENT APPLICATION NUMBER: US/10/325,430
PRIOR FILING DATE: 2002-12-19
PRIOR FILING DATE: 2001-12-19
NUMBER OF SEQ ID NOS: 21
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 11
LENGTH: 6162
TYPE: DNA
ORGANISM: Homo Sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (1)...(6162)
US-10-325-430-11

Query Match Best Local Similarity 91.8%; Score 5661.4; DB 16; Length 6162;

Matches 5944; Conservative 0; Mismatches 16; Indels 243; Gaps 3;

QY 1 ATGTTGAAGTTCAATATATGAGGCGGAAATCTTTGATGCTGCTGCTGCAACCAT 60
DB 1 ATGTTGAAGTTCAATATATGAGGCGGAAATCTTTGATGCTGCTGCTGCAACCAT 60
QY 61 GCACGCGGCGCTCCAGGCTGAATCTGTTCTTCAGAGGAGAAACCAACCTTTATGACTCA 120
DB 61 GCACGCGGCGCTCCAGGCTGAATCTGTTCTTCAGAGGAGAAACCAACCTTTATGACTCA 120
QY 121 CAGCAGATGCTCTCTTTCCCGAGAGGATTTAGATGCTCTCTTTGTTGTTGAA 180
DB 121 CAGCAGATGCTCTCTTTCCCGAGAGGATTTAGATGCTCTCTTTGTTGTTGAA 180
QY 121 CAGCAGATGCTCTCTTTCCCGAGAGGATTTAGATGCTCTCTTTGTTGTTGAA 180
DB 121 CAGCAGATGCTCTCTTTCCCGAGAGGATTTAGATGCTCTCTTTGTTGTTGAA 180
QY 181 GAATGCAAGTCAAGCTGCTCTGATGAAGATTAAAGCACTGAGCAACTTTGTCGGAAGTAT 240
DB 181 GAATGCAAGTCAAGCTGCTCTGATGAAGATTAAAGCACTGAGCAACTTTGTCGGAAGTAT 240
QY 241 TCCGACACCATAGCTGAGTTACAGAGCTCCAGCCTTCGCGAAGAGACTTCGAGTCA 300
DB 241 TCCGACACCATAGCTGAGTTACAGAGCTCCAGCCTTCGCGAAGAGACTTCGAGTCA 300
QY 301 AGCTCTTGAGGTTGCTGCTCACTTTGCTGAAGTGCAGCTGTTAAGAGAAACCAACCGG 360
DB 301 AGCTCTTGAGGTTGCTGCTCACTTTGCTGAAGTGCAGCTGTTAAGAGAAACCAACCGG 360
QY 361 GACATCTATGCTATGAAGATGATGAAGAGGCTTTATGCGCCAGAGAGAGGTTTCA 420
DB 361 GACATCTATGCTATGAAGATGATGAAGAGGCTTTATGCGCCAGAGAGAGGTTTCA 420

QY 421 TTTTGTGAGAGAGCGGAAATATATATCTCGAAGCAAGACCCGTGATCCCCAATTA 480
DB 421 TTTTGTGAGAGAGCGGAAATATATATCTCGAAGCAAGACCCGTGATCCCCAATTA 480
QY 481 CAGTAAGCTTTTGAAGCAAAATATCACTTTATCTGCTATGATATCAAGCTGAGGG 540
DB 481 CAGTAAGCTTTTGAAGCAAAATATCACTTTATCTGCTATGATATCAAGCTGAGGG 540
QY 541 GACTGCTGCTACCTTTTGAATATATGAGGAGCAAGTATGATGATGATGATGATGAT 600
DB 541 GACTGCTGCTACCTTTTGAATATATGAGGAGCAAGTATGATGATGATGATGATGAT 600
QY 601 TACCTAGCTAGCTGATTTTGGCTGTTACAGCCTTCATCTGATGATGATGATGAT 660
DB 601 TACCTAGCTAGCTGATTTTGGCTGTTACAGCCTTCATCTGATGATGATGATGATGAT 660
QY 661 GACATCAAGCTGAGAACTTCTGTTGACCCGACAGGACATCAAGCTGAGGATTTT 720
DB 661 GACATCAAGCTGAGAACTTCTGTTGACCCGACAGGACATCAAGCTGAGGATTTT 720
QY 721 GATCTGCGCGAATAATGAAATTCAGAGATGATGATGATGATGATGATGATGATGAT 780
DB 721 GATCTGCGCGAATAATGAAATTCAGAGATGATGATGATGATGATGATGATGATGAT 780
QY 781 CAGATTTACATGCTCTGAAAGTCTGATGATGATGATGATGATGATGATGATGATGAT 840
DB 781 CAGATTTACATGCTCTGAAAGTCTGATGATGATGATGATGATGATGATGATGATGAT 840
QY 841 GCGCTGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 900
DB 841 GCGCTGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 900
QY 901 TCCCTCTGCGAGAGGAACTCTGCGAGAGACCTTCAATTAATGATTTCCAGCG 960
DB 901 TCCCTCTGCGAGAGGAACTCTGCGAGAGACCTTCAATTAATGATTTCCAGCG 960
QY 961 TTTTGAATTTCCAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1020
DB 961 TTTTGAATTTCCAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1020
QY 1021 TTTTGTGAGGCGCAAG 1080
DB 1021 TTTTGTGAGGCGCAAG 1080
QY 1081 TCTAAATTTGATGAGCAACATTTGATGATGATGATGATGATGATGATGATGATGAT 1140
DB 1081 TCTAAATTTGATGAGCAACATTTGATGATGATGATGATGATGATGATGATGATGAT 1140
QY 1141 TCTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1200
DB 1141 TCTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1200
QY 1201 CCGTGCAGCTGAGGCGCTCAGGCTTCTGAGTGAAGATCCTGTTGTTGAGGTTTTCG 1260
DB 1201 CCGTGCAGCTGAGGCGCTCAGGCTTCTGAGTGAAGATCCTGTTGTTGAGGTTTTCG 1260
QY 1261 TACAGAGAGCACTGAGGATTTCTGATGATGATGATGATGATGATGATGATGATGAT 1320
DB 1261 TACAGAGAGCACTGAGGATTTCTGATGATGATGATGATGATGATGATGATGATGAT 1320
QY 1321 CCGTGCAGCTGAGGCGCTCAGGCTTCTGAGTGAAGATCCTGTTGTTGAGGTTTTCG 1380
DB 1321 CCGTGCAGCTGAGGCGCTCAGGCTTCTGAGTGAAGATCCTGTTGTTGAGGTTTTCG 1380
QY 1381 TCTGAGAGCAAGTGTCAAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1440
DB 1381 TCTGAGAGCAAGTGTCAAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1440
QY 1441 GAGGTGAGGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1500
DB 1441 GAGGTGAGGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1500
QY 1501 TCCCTCTGAG 1560

```
Db 1501 TCCCTCTGAGAGGACCTTGTCTAATCATCAAGAAATGCACTTAAAGCCAAAGT 1560
Qy 1561 TTGAGACAGACAGATGAGAGTGTCCAGAGAGATGACAAAGCACTGACGCTTCTCAT 1620
Db 1561 TTGAGACAGACAGATGAGAGTGTCCAGAGAGATGACAAAGCACTGACGCTTCTCAT 1620
Qy 1621 GATATCAGAGACAGAGCCGGAAGCTCCAAAGAAATCAAAGACAGAGTACAGGCTCAA 1680
Db 1621 GATATCAGAGACAGAGCCGGAAGCTCCAAAGAAATCAAAGACAGAGTACAGGCTCAA 1680
Qy 1681 GTGGAAGAAATGAGTGTATGATGAATCAGTGTGGAAGAGATCTGTCTGACAGAGAGA 1740
Db 1681 GTGGAAGAAATGAGTGTATGATGAATCAGTGTGGAAGAGATCTGTCTGACAGAGAGA 1740
Qy 1741 CGGAGTATCTCTACGATCTGAGCTGAGAGAGTCTGAGCTGTGCTGAGAAATTCAG 1800
Db 1741 CGGAGTATCTCTCTACGATCTGAGCTGAGAGAGTCTGAGCTGTGCTGAGAAATTCAG 1800
Qy 1801 CGGAAACGACAGAAATGTCAGCATTAACCTGTTGAAGCTTAAGATCAAGGGAAGCTGAA 1860
Db 1801 CGGAAACGACAGAAATGTCAGCATTAACCTGTTGAAGCTTAAGATCAAGGGAAGCTGAA 1860
Qy 1861 GTGGAAGAAATGAGTGTGGAATCTGAGAGATCAATGCTGAGACAGGCTCAAAATTCAGAG 1920
Db 1861 GTGGAAGAAATGAGTGTGGAATCTGAGAGATCAATGCTGAGACAGGCTCAAAATTCAGAG 1920
Qy 1921 CTCGAGAGAACTGAGAGAGGCTGTAAAGCCAGACAGAGGCAACGAGCTGCTGAG 1980
Db 1921 CTCGAGAGAACTGAGAGAGGCTGTAAAGCCAGACAGAGGCAACGAGCTGCTGAG 1980
Qy 1981 AATATCCGACGAGCAAGAGAGCGAGCGAGAGGAGCTGAGAGAGCTGAGAGAGCCGAGAG 2040
Db 1945 -----GCAAGAGAGCGAGCGAGAGGAGCTGAGAGAGCTGAGAGAGCCGAGAG 1992
Qy 2041 GATTCTTCTGAGAGCATCAAGAAAGAGCTGTGAGAGCTGAGAGAGCCGCTATTCTCTG 2100
Db 1993 GATTCTTCTGAGAGCATCAAGAAAGAGCTGTGAGAGCTGAGAGAGCCGCTATTCTCTG 2052
Qy 2101 GAGAAACAGGTAAAGAGACTGAGAGAGAGGAGTGAAGAGAAACAGACTGAGAGATGAC 2160
Db 2053 GAGAAACAGGTAAAGAGACTGAGAGAGAGGAGTGAAGAGAAACAGACTGAGAGATGAC 2112
Qy 2161 ATCCAGACAAAAATCCAAACAGATCCAGACAGATGCTGATAAATTCCTGAGCTCGAAGAG 2220
Db 2113 ATCCAGACAAAAATCCAAACAGATCCAGACAGATGCTGATAAATTCCTGAGCTCGAAGAG 2172
Qy 2221 AAAATCTGGGAGGCCCAAGTCTCAGCCAGACACTAGAGTGCACCTGAAACAGAAAGAG 2280
Db 2173 AAAATCTGGGAGGCCCAAGTCTCAGCCAGACACTAGAGTGCACCTGAAACAGAAAGAG 2232
Qy 2281 CAGACATATGAGAAAGATTTAAAGTGTGGAACAATGATTAAGAAAGAAAGCTGTGCTGAC 2340
Db 2233 CAGACATATGAGAAAGATTTAAAGTGTGGAACAATGATTAAGAAAGAAAGCTGTGCTGAC 2292
Qy 2341 AAGGAGACATGAGAAATGATGATGAGACAGAGAGAGAGGCCCATGAGAGAGGAGAA 2400
Db 2293 AAGGAGACATGAGAAATGATGATGAGACAGAGAGAGAGGCCCATGAGAGAGGAGAA 2352
Qy 2401 ATTCTAGCGAGACAGAGAGGAGTATGATGATGATGATGATGATGATGATGATGATGAT 2460
Db 2353 ATTCTAGCGAGACAGAGAGGAGTATGATGATGATGATGATGATGATGATGATGATGAT 2412
Qy 2461 CAGAGATTTGTGAGATCTGTGAGAGCCAAATTAACCTTCCAGCAATAGAGATCTTTTACC 2520
Db 2413 CAGAGATTTGTGAGATCTGTGAGAGCCAAATTAACCTTCCAGCAATAGAGATCTTTTACC 2472
Qy 2521 CAAAGAGAACTGAAGGCGCAAGAGAGATGATTTCTGAACTCAGGCAAGCAAGAAATTTTAC 2580
Db 2473 CAAAGAGAACTGAAGGCGCAAGAGAGATGATTTCTGAACTCAGGCAAGCAAGAAATTTTAC 2532
Qy 2581 CTGAGACACAGAGCTGGAGATGAGAGCCAGAGACCAAGAAACTGAGAGACAGCTGAG 2640
Db 2533 CTGAGACACAGAGCTGGAGATGAGAGCCAGAGACCAAGAAACTGAGAGACAGCTGAG 2592
Qy 2641 AGATTAACCAACCAAGACAGATGACAGAAATCGGCTGTGGAATCTGAGACAGATTTG 2700
Db 2593 AGATTAACCAACCAAGACAGATGACAGAAATCGGCTGTGGAATCTGAGACAGATTTG 2652
Qy 2701 CGGAGAGTCAAGTATGAGACAGAGAGAGAGAACTGGAAGTCAAGGCGCAGCTCAAGAG 2760
Db 2653 CGGAGAGTCAAGTATGAGACAGAGAGAGAGAACTGGAAGTCAAGGCGCAGCTCAAGAG 2712
Qy 2761 CTACAGCTCTCCCTGAGAGAGCGGAGTCAAGATGAGACAGCCCTGAGAGCTGAGACGGCG 2820
Db 2713 CTACAGCTCTCCCTGAGAGAGCGGAGTCAAGATGAGACAGCCCTGAGAGCTGAGACGGCG 2772
Qy 2821 GCCCTGAGAGAGCGGCTTCCAGGCGGAAGACAGAGCTGGAAGAGACCAAGCAAGAGCT 2880
Db 2773 GCCCTGAGAGAGCGGCTTCCAGGCGGAAGACAGAGCTGGAAGAGACCAAGCAAGAGCT 2832
Qy 2881 GAAAGAGAGATCCAGGACCTCAGGACATGAGATGAAATCCAGGCGCAATTTGATGCT 2940
Db 2833 GAAAGAGAGATCCAGGACCTCAGGACATGAGATGAAATCCAGGCGCAATTTGATGCT 2892
Qy 2941 CTTCTGTAAGCTGTATCTGATATCAGAGACCTGAGAGAGAGCTAAACAGCTGACCGAG 3000
Db 2893 CTTCTGTAAGCTGTATCTGATATCAGAGACCTGAGAGAGAGCTAAACAGCTGACCGAG 2952
Qy 3001 GACAGGCTGAATCTCAACCAACCAAACTTCTAATCTGCAACCACTGATGAGCTTCT 3060
Db 2953 GACAGGCTGAATCTCAACCAACCAAACTTCTAATCTGCAACCACTGATGAGCTTCT 3012
Qy 3061 GGGCGCAACGACAGATTTGTACAACTGCGAAGTGAAGTGAACCAATTCGCGCGGAGATC 3120
Db 3013 GGGCGCAACGACAGATTTGTACAACTGCGAAGTGAAGTGAACCAATTCGCGCGGAGATC 3072
Qy 3121 ACGGAAGAGATGAGAGCTTACCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 3180
Db 3073 ACGGAAGAGATGAGAGCTTACCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 3132
Qy 3181 TGCAACATGCTGAGAGACAGATCATGATTTGAGAGGCTTAAACGATGAGCTGTAGAA 3240
Db 3133 TGCAACATGCTGAGAGACAGATCATGATTTGAGAGGCTTAAACGATGAGCTGTAGAA 3192
Qy 3241 AAGAGCGAGTGGAGAGCTGAGAGAGGCTCTGAGTGTAGAGAAATCCAGTTTGAAG 3300
Db 3193 AAGAGCGAGTGGAGAGGCTGAGAGAGGCTCTGAGTGTAGAGAAATCCAGTTTGAAG 3252
Qy 3301 TGTGGGTTGAGAGAGCTGAGAGAGATGCTGAGACACCGAAGAAACAGAGCAGGCGAGGCC 3360
Db 3253 TGTGGGTTGAGAGAGCTGAGAGAGATGCTGAGACACCGAAGAAACAGAGCAGGCGAGGCC 3312
Qy 3361 GATCAGCGATGACCGAGTCTGCGCAGGCTGAGAGGCTGAGAGGCTGAGAGGCTGAGAGGCT 3420
Db 3313 GATCAGCGATGACCGAGTCTGCGCAGGCTGAGAGGCTGAGAGGCTGAGAGGCTGAGAGGCT 3372
Qy 3421 GAGATCTGCTCTGAGAGAGGCTCTCAAGAGAGAGCTGAGAGGCTGAGAGGCTCTCT 3480
Db 3373 GAGATCTGCTCTGAGAGAGGCTCTCAAGAGAGAGCTGAGAGGCTGAGAGGCTCTCT 3432
Qy 3481 GACAGCTCAATGACCTGGAAGAGAGATGCTATGCTTGAATGAAATGCTCCGAGACTTA 3540
Db 3433 GACAGCTCAATGACCTGGAAGAGAGATGCTATGCTTGAATGAAATGCTCCGAGACTTA 3492
Qy 3541 CAGCAGAGCTGAGAGCTGAGAGAGGCTCAACAGAGGCTTCTGGAAGAGCAAGGCCAA 3600
Db 3493 CAGCAGAGCTGAGAGCTGAGAGAGGCTCAACAGAGGCTTCTGGAAGAGCAAGGCCAA 3552
Qy 3601 TTACAGCAGAGATGAGACCTGAGAGAGAAATCAATTTTCCGCTGACTCAAGAGCTGCA 3660
Db 3553 TTACAGCAGAGATGAGACCTGAGAGAGAAATCAATTTTCCGCTGACTCAAGAGCTGCA 3612
Qy 3661 GAAAGCTTGAATGGGCTGATCTTATGAGAGCAAGAAAGAGTGAATTTGAGATCAGCTG 3720
Db 3613 GAAAGCTTGAATGGGCTGATCTTATGAGAGCAAGAAAGAGTGAATTTGAGATCAGCTG 3672
```

Qy	3721	GAATACTTGAAGTTCTCTATTCTCATGAAAGAGTGAATGAAAGGACATATTTCTCA	3780
Db	3673	GAATACTTGAAGTTCTCTATTCTCATGAAAGAGTGAATGAAAGGACATATTTCTCA	3732
Qy	3781	CAATACCAATCTCATGATTTTCTGCAAGCCAAATGCAACCTGCTAAAGAAA--	3838
Db	3733	CAATACCAATCTCATGATTTTCTGCAAGCCAAATGCAACCTGCTAAAGAAAAG	3792
Qy	3839	-----AGTTCTCTGCACTAC-----	3855
Db	3793	GGTTATTATGTCGACGGAAGAGACCTGCTTAAACCAACAGTTCTCTGCACTAC	3852
Qy	3856	AATGACCTGAAGCTGGCCCTTGGAAGAGGAAAGTCCGTGTCAGACTGAGGAAGCC	3915
Db	3853	AATGACCTGAAGCTGGCCCTTGGAAGAGGAAAGTCCGTGTCAGACTGAGGAAGCC	3912
Qy	3916	CTTCAGAAAGACCCGACCTGAGCTCCGGTCCGGCCGGGAGGAAGCTGACCCGCAAGGCA	3975
Db	3913	CTTCAGAAAGACCCGACCTGAGCTCCGGTCCGGCCGGGAGGAAGCTGACCCGCAAGGCA	3972
Qy	3976	ACGGACCAACCCACCCATTCACGCCAGCCACCGGAGGACAGATGCCATGTCGCC	4035
Db	3973	ACGGACCAACCCACCCATTCACGCCAGGACCGGAGGACAGATGCGCATGTCGCC	4032
Qy	4036	ATCGTGCCTGTCGACGAGACCAAGCCAGTGCATGACCTGTCGCCGCCATTCAC	4095
Db	4033	ATCGTGCCTGTCGACGAGACCAAGCCAGTGCATGACCTGTCGCCGCCATTCAC	4092
Qy	4096	CGCAGAAAGAGTCTTCAACTCCAGAGAAATTTAGTCGGCTCTTAAGAAAGCATGCAC	4155
Db	4093	CGCAGAAAGAGTCTTCAACTCCAGAGAAATTTAGTCGGCTCTTAAGAAAGCATGCAC	4152
Qy	4156	CACAATTTCTCAACCGATTCAAGTAGACTGAACATGCGACCAACAAAGTGTCTGT	4215
Db	4153	CACAATTTCTCAACCGATTCAACCTAGACTGAACATGCGACCAACAAAGTGTCTGT	4212
Qy	4216	TGTCGAGTACCGTGCATTTGGAAGGACGATCAAAATGTCGAAATGTCAGATG	4275
Db	4213	TGTCGAGTACCGTGCATTTGGAAGGACGATCAAAATGTCGAAATGTCAGATG	4272
Qy	4276	TGTCACCCCAAGTGTCCAGCTGTTCAGACCACTGCGCTTGCTGTGATATG	4335
Db	4273	TGTCACCCCAAGTGTCCAGCTGTTCAGACCACTGCGCTTGCTGTGATATG	4332
Qy	4336	ACACCTTCAACGAGGCTTCTGCGCTGACAAATGAACTCCAGGCTTCACAGCCAG	4395
Db	4333	ACACCTTCAACGAGGCTTCTGCGCTGACAAATGAACTCCAGGCTTCACAGCCAG	4392
Qy	4396	GAGCCACAGACAGCTTGCACCTGGAAGGGTGAAGAAGTGCCACAGAAATTAACAG	4455
Db	4393	GAGCCACAGACAGCTTGCACCTGGAAGGGTGAAGAAGTGCCACAGAAATTAACAG	4452
Qy	4456	GGAACGACGAGCTGGGACAGGAAGTACATTTGCTGGAAGGATCAAAAGTCTCATTTAT	4515
Db	4453	GGAACGACGAGCTGGGACAGGAAGTACATTTGCTGGAAGGATCAAAAGTCTCATTTAT	4512
Qy	4516	GACATAGAGGACAGAACTGGAACAAGGCTGGTGAAGAAATTTGAGCTGTGCTTCC	4575
Db	4513	GACATAGAGGACAGAACTGGAACAAGGCTGGTGAAGAAATTTGAGCTGTGCTTCC	4572
Qy	4576	GACGGGGATGATCTATTCATTCATGATGCGGTTGATCTCCAACTCGCAATTAAGCCAA	4635
Db	4573	GACGGGGATGATCTATTCATTCATGATGCGGTTGATCTCCAACTCGCAATTAAGCCAA	4632
Qy	4636	GCAAGTGTCCATACATTACTGAAGATGAAATTCACCCGACACCACTGTGCGCCGG	4695
Db	4633	GCA-----	4635
Qy	4696	AGAACCTCTACTTGATGCTCCAGCTTCCCTGACAAACAGGCGTGGGTCAACCGCTTA	4755
Db	4636	-----	4635

QY	4756	GAAATCAAGTTGTGGCAGAGTGGAGAGATTTCATAGGAAAAAGCAGAAAGCTGATGCTAACTG	4815
Db	4636	-----GAAAAAGCAGAAAGCTGAAGTCTAACTG	4662
QY	4816	CTTGAAAACTCCCTGCTGAAACTGGAAAGGTATGACCGTCTAGACATGAATGACACGCTG	4875
Db	4663	CTTGGAAACTCCCTGCTGAAACTGGAAAGGTATGACCGTCTAGACATGAATGACACGCTG	4722
QY	4876	CCCTTCAGTACACAGATGATGTGTGGGACCCGAGAAAGGACTTACGCCCTGATATGTC	4935
Db	4723	CCCTTCAGTACACAGATGATGTGTGGGACCCGAGAAAGGACTTACGCCCTGATATGTC	4782
QY	4936	TTGAAAAAATCCCTTAAACCAATGTCCAGAAATTGAGACATGCTTCCAAATTTATATATC	4995
Db	4783	TTGAAAAAATCCCTTAAACCAATGTCCAGAAATTGAGACATGCTTCCAAATTTATATATC	4842
QY	4996	AAGGACCTGGAGAAACTCTATATATACAGAGAAAGGCGGGACATGTGTCTTGTGAC	5055
Db	4843	AAGGACCTGGAGAAACTCTATATATACAGAGAAAGGCGGGACATGTGTCTTGTGAC	4902
QY	5056	GTTAGAAAAATGTAACAGTATCCCTGGCCAGATCCACATGACCTGTGCCAGCCGACATCTCA	5115
Db	4903	GTTAGAAAAATGTAACAGTATCCCTGGCCAGATCCACATGACCTGTGCCAGCCGACATCTCA	4962
QY	5116	CCCAACATTTTGAAGCTGTGAAGGCGCTGCACATTTGTTGGGCGACGCAAGATTTAGAAC	5175
Db	4963	CCCAACATTTTGAAGCTGTGAAGGCGCTGCACATTTGTTGGGCGACGCAAGATTTAGAAC	5022
QY	5176	GGGCTCTGCATCTGTGACGCAATGCCAGAAATTCGCTTCTCCGCTCAACAGAAAC	5235
Db	5023	GGGCTCTGCATCTGTGACGCAATGCCAGAAATTCGCTTCTCCGCTCAACAGAAAC	5082
QY	5236	CTCAGCAATATCTGCATCCGAAAAGATAGAGACTCAGAGCCCTCAGCTGATATCAC	5295
Db	5083	CTCAGCAATATCTGCATCCGAAAAGATAGAGACTCAGAGCCCTCAGCTGATATCAC	5142
QY	5296	TTACCAATTAACAGTATCTCATTTGGAACCAATTAATTTACAGAAATGACATGAAGCAG	5355
Db	5143	TTACCAATTAACAGTATCTCATTTGGAACCAATTAATTTACAGAAATGACATGAAGCAG	5202
QY	5356	TACAGCTCTGAGGAATTCCTGGATTAAGAAATGACCATTCCTTTGGCACCTGCTGTGTTGCC	5415
Db	5203	TACAGCTCTGAGGAATTCCTGGATTAAGAAATGACCATTCCTTTGGCACCTGCTGTGTTGCC	5262
QY	5416	GCCCTTTCACAACAGCTTCCCTGTCTCAATGATGTGAGGTGAACACGCGAGGGCAGCAGAG	5475
Db	5263	GCCCTTTCACAACAGCTTCCCTGTCTCAATGATGTGAGGTGAACACGCGAGGGCAGCAGAG	5322
QY	5476	GAGTACTGTGTGTTTCCAGAAATTTGAGATGTTGTGTGATTTCTTACGGAAGCCTTAC	5535
Db	5323	GAGTACTGTGTGTTTCCAGAAATTTGAGATGTTGTGTGATTTCTTACGGAAGCCTTAC	5382
QY	5536	CGCACAACGATCTCAAGTGAAGTCGTTACTTTGGCCTTTGCTTACAGAGAACCTTAT	5595
Db	5383	CGCACAACGATCTCAAGTGAAGTCGTTACTTTGGCCTTTGCTTACAGAGAACCTTAT	5442
QY	5596	CTGTTTGGACCACTTCAACTCACTCACTGAAATTAAGATTCAGGCAACGCTCTCAGCA	5655
Db	5443	CTGTTTGGACCACTTCAACTCACTCACTGAAATTAAGATTCAGGCAACGCTCTCAGCA	5502
QY	5656	GGGACCCCTCGCAGCGATCTTGACATCCGAAACCGCGCTACCTGAGGCCCTGCATT	5715
Db	5503	GGGACCCCTCGCAGCGATCTTGACATCCGAAACCGCGCTACCTGAGGCCCTGCATT	5562
QY	5716	TCCTCAGAGCAGATTTACTTGGCGTCTCTATACAGAGATTAATTAAGGTCATTTGCTTCG	5775
Db	5563	TCCTCAGAGCAGATTTACTTGGCGTCTCTATACAGAGATTAATTAAGGTCATTTGCTTCG	5622
QY	5776	AAGGAAACCTCGTGAAGAGATCCGGAATGAAACACACCGGGGCGCTTCAACTTCCGC	5835
Db	5623	AAGGAAACCTCGTGAAGAGATCCGGAATGAAACACACCGGGGCGCTTCAACTTCCGC	5682
QY	5836	AGCAGCCCCCAAGCAGGCGCCACCAAGTCAAGAGCAGTCAACAGAGCAGTCAACAGGCGGTGCGC	5895

Db 5683 AGACGCCCAACAGGAGGAGCCCAACGATCAACAGACGAGTCCGCTGAGCC 5742
Qy 5896 TCCAGGCCAGCGCGCGCCGAGAGCCCAAGCCAGAGCCCAAGCCAGCCAGCCG 5955
Db 5743 TCCAGGCCAGCGCGCGCCGAGAGCCCAAGCCAGAGCCCAAGCCAGCCAGCCG 5802
Qy 5956 TACCGGAGGAGGAGGAGCCGAGGAGTCTGCGGAGGAGTCTGCGGAGGAGGAG 6015
Db 5803 TACCGGAGGAGGAGGAGCCGAGGAGTCTGCGGAGGAGTCTGCGGAGGAGGAG 5862
Qy 6016 GAGAAATCCCCCGCGCGAGTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 6075
Db 5863 GAGAAATCCCCCGCGCGAGTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 5922
Qy 6076 GAGAGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 6135
Db 5923 GAGAGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 5982
Qy 6136 AACAGGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 6158
Db 5983 AACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 6005

RESULT 13
US-10-017-216-1
; Sequence 1, Application US/10017216
; Publication No. US20020160483A1
; GENERAL INFORMATION:
; APPLICANT: KABELLER-LIBERMANN, Rosana
; TITLE OF INVENTION: 13245, A No. US20020160483A1e1 Human Myotonic Dystrophy Type Prod
; FILE REFERENCE: 10147-57U1
; CURRENT APPLICATION NUMBER: US/10/017,216
; PRIOR FILING DATE: 2001-10-23
; PRIOR APPLICATION NUMBER: US 60/242,429
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 1
; LENGTH: 6574
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-017-216-1

Query Match 91.8%; Score 5661.4; DB 13; Length 6574;
Best Local Similarity 95.8%; Pred. No. 0;
Matches 5944; Conservative 0; Mismatches 16; Indels 243; Gaps 3;

Qy 1 ATGTGAAGTTCAAAATATGAGGAGGAGGAGTCTTGGATGCTGATGCTGAAACCAT 60
Db 19 ATGTGAAGTTCAAAATATGAGGAGGAGGAGTCTTGGATGCTGATGCTGAAACCAT 78
Qy 61 GCCAGCGGAGGAGGAGGAGGAGTCTTGGATGCTGATGCTGAAACCAT 120
Db 79 GCCAGCGGAGGAGGAGGAGGAGTCTTGGATGCTGATGCTGAAACCAT 138
Qy 121 CAGCAATGCT 180
Db 139 CAGCAATGCT 198
Qy 181 GAATGAGTCAAGCT 240
Db 199 GAATGAGTCAAGCT 258
Qy 241 TCCGACACCATAGCTAGTATGAGGAGGAGTCTTGGATGCTGATGCTGAAACCAT 300
Db 259 TCCGACACCATAGCTAGTATGAGGAGGAGTCTTGGATGCTGATGCTGAAACCAT 318
Qy 301 AGCTTGTAGT 360
Db 319 AGCTTGTAGT 378

Qy 361 GACATCTATGCTATGAAAGTATGAAAGGAGGAGTCTTATGAGGAGGAGGAGGAGTCTCA 420
Db 379 GACATCTATGCTATGAAAGTATGAAAGGAGGAGTCTTATGAGGAGGAGGAGGAGTCTCA 438
Qy 421 TTTTGTAGGAGGAGGAGGAGGAGTCTTATGAGGAGGAGGAGGAGGAGGAGGAGGAGTCA 480
Db 439 TTTTGTAGGAGGAGGAGGAGGAGTCTTATGAGGAGGAGGAGGAGGAGGAGGAGGAGTCA 498
Qy 481 CAGTATGCTCTTCAAGCAAAATCACTTATGAGGAGGAGGAGGAGGAGGAGGAGGAGTCA 540
Db 499 CAGTATGCTCTTCAAGCAAAATCACTTATGAGGAGGAGGAGGAGGAGGAGGAGGAGTCA 558
Qy 541 GACTTGTCTCACTTTTGAATATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGTCA 600
Db 559 GACTTGTCTCACTTTTGAATATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGTCA 618
Qy 601 TACCTAGCTAGGAGGAGTCTTATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGTCA 660
Db 619 TACCTAGCTAGGAGGAGTCTTATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGTCA 678
Qy 661 GACATCAAGGCTGAGAAATCTCTGTTGACCGGACAGACATCAAGCTGAGTCA 720
Db 679 GACATCAAGGCTGAGAAATCTCTGTTGACCGGACAGACATCAAGCTGAGTCA 728
Qy 721 GATCTGCGCGGAGAAATGAAATTCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGTCA 780
Db 739 GATCTGCGCGGAGAAATGAAATTCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGTCA 798
Qy 781 CCAATTAAGGCTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 840
Db 799 CCAATTAAGGCTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 858
Qy 841 GGCCTGGAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 900
Db 859 GGCCTGGAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 918
Qy 901 TCCCTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 960
Db 919 TCCCTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 978
Qy 961 TTTTGAATTTTCAAGTATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1020
Db 979 TTTTGAATTTTCAAGTATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1038
Qy 1021 TTGTTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1080
Db 1039 TTGTTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1098
Qy 1081 TCTAAATGATGAGAAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1140
Db 1099 TCTAAATGATGAGAAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1158
Qy 1141 TCTGAGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1200
Db 1159 TCTGAGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1218
Qy 1201 CCGTGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1260
Db 1219 CCGTGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1278
Qy 1261 TACAGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1320
Db 1279 TACAGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1338
Qy 1321 CCGTGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1380
Db 1339 CCGTGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1398
Qy 1381 TCTCAAGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1440
Db 1399 TCTCAAGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1458
Qy 1441 GAGTGAAGGCTGTGCTTATGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1500

1459 GAGGTGAGGCTGTGCTTTAGTCAGAGAGAGGTGAGGCTGAGGCTCTTGAGACTTCAGAGA 1518
1501 TCCCTCTTGAGACAGGACTTTGCTTACCTTACATCA CAGAAATGCACTAGCTTTAAACGAACT 1560
1519 TCCCTCTTGAGACAGGACTTTGCTTACCTTACATCA CAGAAATGCACTAGCTTTAAACGAACT 1578
1561 TTGGAGCAACACGAGATGAGAGGTGCTCCAGAGAGATGACAAAGACTGCACTTCTTCAT 1620
1579 TTGGAGCAACACGAGATGAGAGGTGCTCCAGAGAGATGACAAAGACTGCACTTCTTCAT 1638
1621 GATATCAGAGACAGAGACCCGAGGCTCCAGAAATCAAAGAGCAGAGTACAGGCTCAA 1680
1639 GATATCAGAGACAGAGACCCGAGGCTCCAGAAATCAAAGAGCAGAGTACAGGCTCAA 1698
1681 GTGGAGAAATGAGGTTGATGATGAATCACTTGGAGAGAGATCTTGTCTTCAGCAAGAGA 1740
1699 GTGGAGAAATGAGGTTGATGATGAATCACTTGGAGAGAGATCTTGTCTTCAGCAAGAGA 1758
1741 CGGAGTGAATCTTACGAATCTGAGCTGAGAGAGTCTTGCGCTTGTCTGTAAGAAATTCAG 1800
1759 CGGAGTGAATCTTACGAATCTGAGCTGAGAGAGTCTTGCGCTTGTCTGTAAGAAATTCAG 1818
1801 CGGAAAGCCAGAGATGTCAGCATTAACCTGTTGAAGGCTAAAGATCAAGGGAAGCCTGA 1860
1819 CGGAAAGCCAGAGATGTCAGCATTAACCTGTTGAAGGCTAAAGATCAAGGGAAGCCTGA 1878
1861 GTGGAGAAATATGCGAACTGGAAGATCAATGCTGAGCAGCAGCTCAAAATTCAGAG 1920
1879 GTGGAGAAATATGCGAACTGGAAGATCAATGCTGAGCAGCAGCTCAAAATTCAGAG 1938
1921 CTCGAGAGAAATGAGAGAGGCTGTAAAACGACGAGAGCCACCGAGCTGCTCAG 1980
1939 CTCGAGAGAAATGAGAGAGGCT----- 1962
1981 AATATCCGCGCAAGAGCGAGCGGAGGAGCTGGAGAAAGCTGCAAGACGAGAG 2040
1963 -----GCAAGAGCGAGCGCGAGGAGGAGCTGGAGAAAGCTGCAAGACGAGAG 2010
2041 GATTCCTTGAAAGGATCAGAAAAGAACTGTGGAAGCTGAGGAGCGCGCATTTCTCTG 2100
2011 GATTCCTTGAAAGGATCAGAAAAGAACTGTGGAAGCTGAGGAGCGCGCATTTCTCTG 2070
2101 GAGAACAAAGTAAAGAGCTAGAGACATGAGACCTGAGAGAAAAGCACTGAAGATGAC 2160
2071 GAGAACAAAGTAAAGAGCTAGAGACATGAGACCTGAGAGAAAAGCACTGAAGATGAC 2130
2161 ATCCAGACAAAAATCCCAACAGATCCAGAGATGAGCTGATTAATTTGAGGCTCGAAGAG 2220
2131 ATCCAGACAAAAATCCCAACAGATCCAGAGATGAGCTGATTAATTTGAGGCTCGAAGAG 2190
2221 AAACATCGGAGGCGCCAACTCTCAAGCCAGCACTTGAAGTGCACCTGAAACAGAAAGAG 2280
2191 AAACATCGGAGGCGCCAACTCTCAAGCCAGCACTTGAAGTGCACCTGAAACAGAAAGAG 2250
2281 CAGCATATGAGAGAAAAGATTAAAGTTGGACATCAATCAGATTAAGAAAGACTTGCTGAC 2340
2251 CAGCATATGAGAGAAAAGATTAAAGTTGGACATCAATCAGATTAAGAAAGACTTGCTGAC 2310
2341 AAGAGACACTGAGAGAACTGATGACAGAGACAGAGAGAGAGGCCATGAGAAAGGCGAA 2400
2311 AAGAGACACTGAGAGAACTGATGACAGAGACAGAGAGAGAGGCCATGAGAAAGGCGAA 2370
2401 ATTCTCAGAGAACAGAAAGCGATGATCAATGCTATGATTTCCAAATCAGATCCCTGAA 2460
2371 ATTCTCAGAGAACAGAAAGCGATGATCAATGCTATGATTTCCAAATCAGATCCCTGAA 2430
2461 CAGAGGATTTGAGAACTGTCTGAGCGCAATTAACCTTGACAAATAGCAGCTTTTAAAC 2520
2431 CAGAGGATTTGAGAACTGTCTGAGCGCAATTAACCTTGACAAATAGCAGCTTTTAAAC 2490
2521 CAAAGGAACATGAAAGGCCAAGAAAGATGATTTTGAACCTCAGGCAACAGAAATTTTAC 2580

2491 CAAAGGAACATGAAGCCCAAGAAAGAGATGATTTCTGAACCTCAGGCAACAGAAATTTTAC 2550
2581 CTGAGACACAGGCTGGAGATTGAGAGCCAGAACCGAAACTGAGAGACAGCTGAG 2640
2551 CTGAGACACAGGCTGGAGATTGAGAGCCCAAGAACCGAAACTGAGAGACAGCTGAG 2610
2641 AAGATCAGCCACCAAGACCAAGTGAACAGAAATCGGCTGTGGAACCTGAGACAGATTG 2700
2611 AAGATCAGCCACCAAGACCAAGTGAACAGAAATCGGCTGTGGAACCTGAGACAGATTG 2670
2701 CGGAGGATCAGTTAAGACAGAGAGCAAGAACTGAGAGCTCAAGGCCCTCACAGAG 2760
2671 CGGAGGATCAGTTAAGACAGAGAGCAAGAACTGAGAGCTCAAGGCCCTCACAGAG 2730
2761 CTACAGCTCTCCCTGAGAGAGCGGAGTCAAGTTGACAGCCCTGAGAGCTGACGCGCG 2820
2731 CTACAGCTCTCCCTGAGAGAGCGGAGTCAAGTTGACAGCCCTGAGAGCTGACGCGCG 2790
2821 GCCCTGAGAGCCAGCTTGGCCAGGCGGAAGACAGAGCTGGAAGAGACCAAGAGCT 2880
2791 GCCCTGAGAGCCAGCTTGGCCAGGCGGAAGACAGAGCTGGAAGAGACCAAGAGCT 2850
2881 GAAAGAGAGATCCAGGCACTCAAGGCAATGAGATGAATCCAGGCGAAATTTGATGCT 2940
2851 GAAAGAGAGATCCAGGCACTCAAGGCAATGAGATGAATCCAGGCGAAATTTGATGCT 2910
2941 CTTCTGAACAGCTGTACTGTAATCA CAGACTGAGAGACAGCTTAAACAGCTGACCGAG 3000
2911 CTTCTGAACAGCTGTACTGTAATCA CAGACTGAGAGACAGCTTAAACAGCTGACCGAG 2970
3001 GACAAAGCTGAACTCAACCAAGAACTTTACTTGTCCAAACAACTCGATGAGCTTCT 3060
2971 GACAAAGCTGAACTCAACCAAGAACTTTACTTGTCCAAACAACTCGATGAGCTTCT 3030
3061 GGGCGCAACGAGAGATTGTACAATGCGAGATGAAGTGAACATCTCCGCGGAGATC 3120
3031 GGGCGCAACGAGAGATTGTACAATGCGAGATGAAGTGAACATCTCCGCGGAGATC 3090
3121 ACGGAACGAGAGATCAGCTTACAGCCAGAGCAAAACGATGAGGCTTGAAGACAG 3180
3091 ACGGAACGAGAGATCAGCTTACAGCCAGAGCAAAACGATGAGGCTTGAAGACAG 3150
3181 TGCACATGCTGAGAGAAACAGTTCATGATTTGAGAGCCCTTAAACGATGAGCTGTAGAA 3240
3151 TGCACATGCTGAGAGAAACAGTTCATGATTTGAGAGCCCTTAAACGATGAGCTGTAGAA 3210
3241 AAAGAGCGGAGTGGAGGCGCTGAGAGAGCGCTCGGGGTGATGAGAAATCCAGTTTGA 3300
3211 AAAGAGCGGAGTGGAGGCGCTGAGAGAGCGCTCGGGGTGATGAGAAATCCAGTTTGA 3270
3301 TGTGGGTTTGAAGACTCAGAGATGCTGAGACCCGAGAAACAGAGAGGCGAGAGCC 3360
3271 TGTGGGTTTGAAGACTCAGAGATGCTGAGACCCGAGAAACAGAGAGGCGAGAGCC 3330
3361 GATCAGCGGATCAACGAGTCTCGCAGGTTGAGAGCTGAGCATGTAAGAGCAAGGCT 3420
3331 GATCAGCGGATCAACGAGTCTCGCAGGTTGAGAGCTGAGCATGTAAGAGCAAGGCT 3390
3421 GAGATTCGCGCTGAGACAGGCTCTCAAGAGAGAGCTGAAGCGCGAGAGCTTCT 3480
3391 GAGATTCGCGCTGAGACAGGCTCTCAAGAGAGAGCTGAAGCGCGAGAGCTTCT 3450
3481 GACAAAGCTCAATGACCTTGAAGAGAGCATGTATGCTTGAATGAATGCCCCGAAGCTTA 3540
3451 GACAAAGCTCAATGACCTTGAAGAGAGCATGTATGCTTGAATGAATGCCCCGAAGCTTA 3510
3541 CAGCAAGAGCTGAGAGCTGAACGAGAGCTCAACAGAGGCTTCTGGAAGCAAGCCAAA 3600
3511 CAGCAAGAGCTGAGAGCTGAACGAGAGCTCAACAGAGGCTTCTGGAAGCAAGCCAAA 3570
3601 TTACAGCAGCAGATGAGACTGCAAGAAATCAATTTTCCGTCTGACTCAAGAGACTGCA 3660
3571 TTACAGCAGCAGATGAGACTGCAAGAAATCAATTTTCCGTCTGACTCAAGAGACTGCA 3630

Db 5641 AAGGAAACCTCGTAGAGAGTCCGGCACTGAACACACCGGGCCCGTCCACTTCCCGC 5700
Qy 5836 AGAGCCCAACAAAGCGGCCCAACCGTACAGACACATACCAAGCGGTGGCC 5895
Db 5701 AGAGCCCAACAAAGCGGCCCAACCGTACAGACACATACCAAGCGGTGGCC 5760
Qy 5896 TCCAGCCCAACCGGCCCAAGCGGCCCAACCGTACAGACACATACCAAGCGGTGGCC 5955
Db 5761 TCCAGCCCAACCGGCCCAAGCGGCCCAACCGTACAGACACATACCAAGCGGTGGCC 5820
Qy 5956 TCCAGCCCAACCGGCCCAAGCGGCCCAACCGTACAGACACATACCAAGCGGTGGCC 6015
Db 5821 TCCAGCCCAACCGGCCCAAGCGGCCCAACCGTACAGACACATACCAAGCGGTGGCC 5880
Qy 6016 GAGAGTCCCGCGCGATGATCTACAGACCGGAGAGAGCGGTCCCGCGAGGTGTTT 6075
Db 5881 GAGAGTCCCGCGCGATGATCTACAGACCGGAGAGAGCGGTCCCGCGAGGTGTTT 5940
Qy 6076 GAGAGTCCCGCGCGATGATCTACAGACCGGAGAGAGCGGTCCCGCGAGGTGTTT 6135
Db 5941 GAGAGTCCCGCGCGATGATCTACAGACCGGAGAGAGCGGTCCCGCGAGGTGTTT 6000
Qy 6136 AACAGGTCTGGAGCAGTCTTC 6158
Db 6001 AACAGGTCTGGAGCAGTCTTC 6023

RESULT 14
US-10-325-430-10
; Sequence 10, Application US/10325430
; Publication No. US20030153525A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc
; APPLICANT: Silos-Santiago, Immaculada
; APPLICANT: Rosenfeld, Julie Beth
; TITLE OF INVENTION: METHODS AND COMPOSITIONS IN TREATING
; TITLE OF INVENTION: PAIN AND PAINFUL DISORDERS USING 1465, 1587, 2146, 2207,
; TITLE OF INVENTION: 32838, 336 and 52908
; FILE REFERENCE: MP101-294PRM
; CURRENT APPLICATION NUMBER: US/10325,430
; PRIOR FILING DATE: 2002-12-19
; PRIOR APPLICATION NUMBER: US 60/341,953
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 6574
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-10-325-430-10

Query Match 91.8%; Score 5661.4; DB 16; Length 6574;
Best Local Similarity 95.8%; Pred. No. 0;
Matches 5944; Conservative 0; Mismatches 16; Indels 243; Gaps 3;

Qy 1 ATGTGAACTTCAATATGAGCGCGGAATCCCTTGGATGCTGTGCTGAACCCATT 60
Db 19 ATGTGAACTTCAATATGAGCGCGGAATCCCTTGGATGCTGTGCTGAACCCATT 78
Qy 61 GCCACCGCGGCTCCAGGCTGAATCTGTCTTCCAGGGGAAACACCCCTTATGACTCA 120
Db 79 GCCACCGCGGCTCCAGGCTGAATCTGTCTTCCAGGGGAAACACCCCTTATGACTCA 138
Qy 121 CAGCAGATGTCTCTCTTCCGAGAGAGGATATAGATGCGCTCTTGTCTCTTTTGA 180
Db 139 CAGCAGATGTCTCTCTTCCGAGAGAGGATATAGATGCGCTCTTGTCTCTTTTGA 198
Qy 181 GAATCAAGTCAAGCTCTCTGATGAAGATTAAGCACTGAGCACTTTGTCCGGAAGTAT 240
Db 199 GAATCAAGTCAAGCTCTCTGATGAAGATTAAGCACTGAGCACTTTGTCCGGAAGTAT 258
Qy 241 TCCGACACCATAGCTGATGATGAGAGCTCCAGCTTGGGCAAGACCTTGCAGAGTCA 300

Db 259 TCCGACACCATAGCTGATGATGAGAGCTCCAGCTTCCGCAAGACCTTGCAGAGTCA 318
Qy 301 AGCTTGTAGGTGTGTGCTCACTTGTGTAAGTCAAGTGTGTAAGAGAGAAACCGGG 360
Db 319 AGCTTGTAGGTGTGTGCTCACTTGTGTAAGTCAAGTGTGTAAGAGAGAAACCGGG 378
Qy 361 GACATCATATGTAAGATGATGAAGAGAGGCTTTATTTGGCCAGAGAGGTTTCA 420
Db 379 GACATCATATGTAAGATGATGAAGAGAGGCTTTATTTGGCCAGAGAGGTTTCA 438
Qy 421 TTTTGTAGAGAGAGGAGAAATATATCTCCAGAGACAAAGCCCGTGAATCCCCAATTA 480
Db 439 TTTTGTAGAGAGAGGAGAAATATATCTCCAGAGACAAAGCCCGTGAATCCCCAATTA 498
Qy 481 CAGTATGCTTTTCAGAGACAAATATCACTTTATCTGATGATGATATGAGTGAAGG 540
Db 499 CAGTATGCTTTTCAGAGACAAATATCACTTTATCTGATGATGATATGAGTGAAGG 558
Qy 541 GACTTCTGTCACTTTTGAATGATATGAGACCAAGTATGATGATGATGATGATGAT 600
Db 559 GACTTCTGTCACTTTTGAATGATATGAGACCAAGTATGATGATGATGATGATGAT 618
Qy 601 TACCTTGTAGAGTGAATTTTGTGCTTCAAGGCTTCACTGATGATGATGATGATGAT 660
Db 619 TACCTTGTAGAGTGAATTTTGTGCTTCAAGGCTTCACTGATGATGATGATGATGAT 678
Qy 661 GACATCAAGCTTTCAGAGACATTTCTGTTGACCGCACAGACATCAAGCTGTGATTTT 720
Db 679 GACATCAAGCTTTCAGAGACATTTCTGTTGACCGCACAGACATCAAGCTGTGATTTT 738
Qy 721 GATCTGCGCGCAAAATGATTAACCAAGATGATGATGATGATGATGATGATGATGAT 780
Db 739 GATCTGCGCGCAAAATGATTAACCAAGATGATGATGATGATGATGATGATGATGAT 798
Qy 781 CAGATTAATAGGCTTCCAGAGTGAATGATGATGATGATGATGATGATGATGATGAT 840
Db 799 CAGATTAATAGGCTTCCAGAGTGAATGATGATGATGATGATGATGATGATGATGAT 858
Qy 841 GGCCTGAGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 900
Db 859 GGCCTGAGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 918
Qy 901 TCCCTCTTTCAGAGAGGAACTTCTCCAGAACCTTCAATTAATGATGATGATGATGAT 960
Db 919 TCCCTCTTTCAGAGAGGAACTTCTCCAGAACCTTCAATTAATGATGATGATGATGAT 978
Qy 961 TTTTGTAAATTTCCAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1020
Db 979 TTTTGTAAATTTCCAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1038
Qy 1021 TTTTGTGCGGCGAGAGAGAGCTGAAGTTTGAAGTCTTGTGCGATCCTTCTTCTTC 1080
Db 1039 TTTTGTGCGGCGAGAGAGAGCTGAAGTTTGAAGTCTTGTGCGATCCTTCTTCTTC 1098
Qy 1081 TCTAAATTTGACTGGAACAACTTGTGTAATCTCTCCCTTGTGTTCCACCTCTCAAG 1140
Db 1099 TCTAAATTTGACTGGAACAACTTGTGTAATCTCTCTCCCTTGTGTTCCACCTCTCAAG 1158
Qy 1141 TCTGACATGACACTTCCAAATTTTGAATGATGATGATGATGATGATGATGATGATGAT 1200
Db 1159 TCTGACATGACACTTCCAAATTTTGAATGATGATGATGATGATGATGATGATGATGAT 1218
Qy 1201 CCGTCCAGCTGAGGCGCTTCCAGGCTTCCGAGTGAAGAACTGCTTGTGGGGTTTTCG 1260
Db 1219 CCGTCCAGCTGAGGCGCTTCCAGGCTTCCGAGTGAAGAACTGCTTGTGGGGTTTTCG 1278
Qy 1261 TACAGCAAGGCACTGGGGATTTCTGTGATGATCTGAGTCTGTGTCTGGATCTGAGACTCC 1320
Db 1279 TACAGCAAGGCACTGGGGATTTCTGTGATGATCTGAGTCTGTGTCTGGATCTGAGACTCC 1338
Qy 1321 CTTGCGAAGCTAGCTTCAATGAGAAAGAACTTCTCATCAAAAGCAAGAGCTACAGAC 1380

Db 1339 CTTGCCAGACTAGCTTCATGAAAAGAACTTCTCATCAAAAACAAAGACTCAAGAC'1398
Qy 1381 TCTCAGGACAAAGTGTCACAAGATGAGCAGAAAATGACCCGGTTACATCCGAGAGTGTCA 1440
Db 1399 TCTCAGGACAAAGTGTCACAAGATGAGCAGAAAATGACCCGGTTACATCCGAGAGTGTCA 1458
Qy 1441 GAGGTGAGGCTGTGCTTAACTCAGAGGAGGTGAGCTGAAAGCTCTTGAGACTCAGAGA 1500
Db 1459 GAGGTGAGGCTGTGCTTAACTCAGAGGAGGTGAGCTGAAAGCTCTTGAGACTCAGAGA 1518
Qy 1501 TCCCTCTGAGAGAGCCTTGCTACCTACATCAACAATGCACTACTTAAAGCGAAGT 1560
Db 1519 TCCCTCTGAGAGAGCCTTGCTACCTACATCAACAATGCACTACTTAAAGCGAAGT 1578
Qy 1561 TTGAGACAAACGAGATGAGAGTGTCCAGAGAGATGACAAAGCACTGCAGCTTCTCAT 1620
Db 1579 TTGAGACAAACGAGATGAGAGTGTCCAGAGAGATGACAAAGCACTGCAGCTTCTCAT 1638
Qy 1621 GATATCAGAGAGCAGAGCCGGAAGCTCAAGAAATCAAGAGCAGAGTACCAAGCTCAA 1680
Db 1639 GATATCAGAGAGCAGAGCCGGAAGCTCAAGAAATCAAGAGCAGAGTACCAAGCTCAA 1698
Qy 1681 GTGGAAGAAATGAGTGTGATGATGAAATCAGTTGGAAGAGATCTTGTCTCAGCAAGAGA 1740
Db 1699 GTGGAAGAAATGAGTGTGATGATGAAATCAGTTGGAAGAGATCTTGTCTCAGCAAGAGA 1758
Qy 1741 CGGAGTATCTTACGAAATCTGAGCTGAGAGAGTCTCGGCTTGCTGCTGAAGAAATTCAG 1800
Db 1759 CGGAGTATCTTACGAAATCTGAGCTGAGAGAGTCTCGGCTTGCTGCTGAAGAAATTCAG 1818
Qy 1801 CGGAAACGACAGAAATGTCAGCAATTAACCTGTAAGGCTTAAGATCAAGGAAAGCTGAA 1860
Db 1819 CGGAAACGACAGAAATGTCAGCAATTAACCTGTAAGGCTTAAGATCAAGGAAAGCTGAA 1878
Qy 1861 GTGGAAGAAATGTCGAAACTGAGAGAGATCAATGCTGAGCAGAGCTCAAAATTCAGAG 1920
Db 1879 GTGGAAGAAATGTCGAAACTGAGAGAGATCAATGCTGAGCAGAGCTCAAAATTCAGAG 1938
Qy 1921 CTCGAAGAGAACTGAGAGAGGCTGTAAAGCCAGCAGAGGCGCACGAGCTGCTCAG'1980
Db 1939 CTCGAAGAGAACTGAGAGAGGCT----- 1962
Qy 1981 AATATCCGACAGAGAGAGAGCCGAGAGGAGCTGAGAGAGCTGCAAGAACGAGAG 2040
Db 1963 -----GCAAGGAGCCAGCCGAGAGGAGCTGAGAGAGCTGCAAGAACGAGAG 2010
Qy 2041 GATTCTTCTGAGAGCATCAAGAAAGCTGAGAGCTGAGAGAGCCGCCCATTTCTCTG 2100
Db 2011 GATTCTTCTGAGAGCATCAAGAAAGCTGAGAGCTGAGAGAGCCGCCCATTTCTCTG 2070
Qy 2101 GAGAACAAAGGTAAAGACCTAGAGACCATGAGAGCTGAGAGAAACAGACTGAAGATGAC 2160
Db 2071 GAGAACAAAGGTAAAGACCTAGAGACCATGAGAGCTGAGAGAAACAGACTGAAGATGAC 2130
Qy 2161 ATCCAGACAAAAATCCCAACAGATCAGACAGATGAGCTGATAAATTTCTGAGCTGAGAG 2220
Db 2131 ATCCAGACAAAAATCCCAACAGATCAGACAGATGAGCTGATAAATTTCTGAGCTGAGAG 2190
Qy 2221 AAACATCGGAGGCGCCAAAGTCTCAGCCACGACCTAGAGTGCACCTGAAACAGAAAGAG 2280
Db 2191 AAACATCGGAGGCGCCAAAGTCTCAGCCACGACCTAGAGTGCACCTGAAACAGAAAGAG 2250
Qy 2281 CAGACATATGAGAGAAAGATTTAAAGTGTGAGCAATAGATTAAGAAAGAAAGCTGAGTAC 2340
Db 2251 CAGACATATGAGAGAAAGATTTAAAGTGTGAGCAATAGATTAAGAAAGAAAGCTGAGTAC 2310
Qy 2341 AAGGAGACATGAGAGACATGATGACAGACACGAGAGAGAGGCGCCATGAGAGGCGCAA 2400
Db 2311 AAGGAGACATGAGAGACATGATGACAGACACGAGAGAGAGGCGCCATGAGAGGCGCAA 2370
Qy 2401 ATTCTCAGCGAACAAGAGCGATGATCAATGCTATGATTCCAAAGATCAGATCCCTGAA 2460
Db 2371 ATTCTCAGCGAACAAGAGCGATGATCAATGCTATGATTCCAAAGATCAGATCCCTGAA 2430

Qy 2461 CAGAGATGTGTGAATGTGTCTGAGCCCAATTAACCTTGACCAAAATAGCAATCTTTTACC 2520
Db 2431 CAGAGATGTGTGAATGTGTCTGAGCCCAATTAACCTTGACCAAAATAGCAATCTTTTACC 2490
Qy 2521 CAAAGGACATGAAAGCCCAAGAGATGATTTCTGAATCTCAGGCGCAAGAAATTTTAC 2580
Db 2491 CAAAGGACATGAAAGCCCAAGAGATGATTTCTGAATCTCAGGCGCAAGAAATTTTAC 2550
Qy 2581 CTGAGACACAGGCTGGAGAGTTGAGAGCCAGAAACGAAATCTGAGAGCAGCTGAG 2640
Db 2551 CTGAGACACAGGCTGGAGAGTTGAGAGCCAGAAACGAAATCTGAGAGCAGCTGAG 2610
Qy 2641 AAGATCAGCCACCAAGACCAACAGTGAACAAGATCGGCTGCTGAACTGAGACAAAGTTG 2700
Db 2611 AAGATCAGCCACCAAGACCAACAGTGAACAAGATCGGCTGCTGAACTGAGACAAAGTTG 2670
Qy 2701 CGGAGGTCAGTCTTAAGACACAGAGACAGAACTGAGACTCAAGCCCACTCAGAG 2760
Db 2671 CGGAGGTCAGTCTTAAGACACAGAGACAGAACTGAGACTCAAGCCCACTCAGAG 2730
Qy 2761 CTACAGCTCTCCCTGAGAGCGGAGTCAAGTTGACAGCTCCGACAGCTGCACGGGCG 2820
Db 2731 CTACAGCTCTCCCTGAGAGCGGAGTCAAGTTGACAGCTCCGACAGCTGCACGGGCG 2790
Qy 2821 GCCCTGAGAGCCAGCTTCGCGCAGGCGAAGACAGAGCTGGAAGAGACACAGCAGAACT 2880
Db 2791 GCCCTGAGAGCCAGCTTCGCGCAGGCGAAGACAGAGCTGGAAGAGACACAGCAGAACT 2850
Qy 2881 GAAGAGAGATCCAGGACCTCAGCGCATAGAGTGAATTCAGCGCAATTTGATGCT 2940
Db 2851 GAAGAGAGATCCAGGACCTCAGCGCATAGAGTGAATTCAGCGCAATTTGATGCT 2910
Qy 2941 CTTGTAACAGCTGTACTGTATCAGAGACTGAGAGAGCAGCTTAACAGCTGACCGAG 3000
Db 2911 CTTGTAACAGCTGTACTGTATCAGAGACTGAGAGAGCAGCTTAACAGCTGACCGAG 2970
Qy 3001 GACAAAGCTGAACCTCAACCAACCAAACTTCTAATTGCTCAAAACAATCGATGAGCTTCT 3060
Db 2971 GACAAAGCTGAACCTCAACCAACCAAACTTCTAATTGCTCAAAACAATCGATGAGCTTCT 3030
Qy 3061 GCGCGCAACGACGAGATTTGAACACTGCGAAGTGAAGTGAACCATCTCCGCCGAGATC 3120
Db 3031 GCGCGCAACGACGAGATTTGAACACTGCGAAGTGAAGTGAACCATCTCCGCCGAGATC 3090
Qy 3121 ACGGAACGAGATGACAGCTTACACGCAAGCAACGATGAGGCTGGAAGCAGC 3180
Db 3091 ACGGAACGAGATGACAGCTTACACGCAAGCAACGATGAGGCTGGAAGCAGC 3150
Qy 3181 TGCACCATGCTGAGAGAACAGTCAATGATTTGAGGCGCTTAACGATGAGCTGTAGAA 3240
Db 3151 TGCACCATGCTGAGAGAACAGTCAATGATTTGAGGCGCTTAAACGATGAGCTGTAGAA 3210
Qy 3241 AAAGACCGGAGTGGAGAGCTTGAAGAGCGTCTTGAGTGAAGAAATCCAGTTTGAAG 3300
Db 3211 AAAGACCGGAGTGGAGAGCGTCTGAAGAGCGTCTTGAGTGAAGAAATCCAGTTTGAAG 3270
Qy 3301 TGTGCGGTTGAGAGAGCTGCAAGAGATGCTGGAACCCGAAACAGACAGGCGGAGAGCC 3360
Db 3271 TGTGCGGTTGAGAGAGCTGCAAGAGATGCTGGAACCCGAAACAGACAGGCGGAGAGCC 3330
Qy 3361 GATCAGCGATCACCAGTCTCCGACAGTGTGAGAGCTGCAATGAAAGAGCAAGAGCT 3420
Db 3331 GATCAGCGATCACCAGTCTCCGACAGTGTGAGAGCTGCAATGAAAGAGCAAGAGCT 3390
Qy 3421 GAGATTCGCTCTGAGAGAGGCTCTCAAGACAGAACTGAAGCCGAGAGCTCTCT 3480
Db 3391 GAGATTCGCTCTGAGAGAGGCTCTCAAGACAGAACTGAAGCCGAGAGCTCTCT 3450
Qy 3481 GACAGCTCAATGACCTGAGAGAGCAGTGAAGCTTGAATGATGCGGCAAGCTTA 3540
Db 3451 GACAGCTCAATGACCTGAGAGAGCAGTGAAGCTTGAATGATGCGGCAAGCTTA 3510

QY	354	CAGCAGAACTGCGAGACTGAACGAGAGCTCAACAGAGGCTTCTGAAAGACGACCCAA	3600
Db	3511	CAGCAAAACCTGAGACTGAAGAGAGCTCAACAGAGGCTTCTGAAAGACGACCCAA	3570
QY	3601	TTACAGCACAGATGAGCTCGAGAAAAATGCAATTTCCGCTGACCTCAAGACCTGAA	3660
Db	3571	TTACAGCACAGATGAGCTCGAGAAAAATGCAATTTCCGCTGACCTCAAGACCTGAA	3630
QY	3661	GAAAGCTTAGATCGGGCTGATCTAATGAAAGAGAAAGTGACTTGGAGTATCACTG	3720
Db	3631	GAAAGCTTAGATCGGGCTGATCTAATGAAAGAGAAAGTGACTTGGAGTATCACTG	3690
QY	3721	GAAGAACATTCAGAGTTCTCTAATCTCATGAAAAAGTGAAGTGAAGGCACTATTCTCA	3780
Db	3691	GAAGAACATTCAGAGTTCTCTAATCTCATGAAAAAGTGAAGTGAAGGCACTATTCTCA	3750
QY	3781	CAAAACCAACTGATTGATTTCTGCAAGCAAAATGGAACAACTGCTAAAGAAAAA--	3838
Db	3751	CAAAACCAACTGATTGATTTCTGCAAGCAAAATGGAACAACTGCTAAAGAAAAA	3810
QY	3839	-----AGGTTCTCTGCAAGTAC	3855
Db	3811	GGTTATTATTAGTCGACGGAAGAGAGACCTGCTTACCAACAGAGTTCTCTGCAAGTAC	3870
QY	3856	AATGAGCTGGAAGCTGGCCCTGAGAGAGAGAAAGTCCGCTGTGACAGCTAGAGAGAAC	3915
Db	3871	AATGAGCTGGAAGCTGGCCCTGAGAGAGAGAAAGTCCGCTGTGACAGCTAGAGAGAAC	3930
QY	3916	CTTGAGAAAGACCCGGATGAGGTCGGGTCGGCCGGGGAGAAAGTCCCAACCGAGAAACA	3975
Db	3931	CTTGAGAAAGACCCGGATGAGGTCGGGTCGGCCGGGGAGAAAGTCCCAACCGAGAAACA	3990
QY	3976	ACGAGCAACCCACACCCATTCACCGCCAGCCACCGCGAGGACAGAGATCGCATGTCCGC	4035
Db	3991	ACGAGCAACCCACACCCATTCACCGCCAGCCACCGCGAGGACAGAGATCGCATGTCCGC	4050
QY	4036	ATCGTCGGGTGCGCAGAGACCAAGCCCATGTGCATGAGCTTGCGCCCGCCATCCAGC	4095
Db	4051	ATCGTCGGGTGCGCAGAGACCAAGCCCATGTGCATGAGCTTGCGCCCGCCATCCAGC	4110
QY	4096	CGCAGAAAGAGGCTTCAACTCCAGAGGAATTTAGTGGCGCTCTAAGGAAACGATGCAC	4155
Db	4111	CGCAGAAAGAGGCTTCAACTCCAGAGGAATTTAGTGGCGCTCTAAGGAAACGATGCAC	4170
QY	4156	CACAAATTTCTCTACCGAATTCAGATGAGTGAAGCTGAACATGCGACCAAAAGTGTCTGTG	4215
Db	4171	CACAAATTTCTCTACCGAATTCAGATGAGTGAAGCTGAACATGCGACCAAAAGTGTCTGTG	4230
QY	4216	TGTCTGATACCGGTGCACCTTTGGAACCGCAGGCATCCAAATGTCTGAAATGCAGGTATG	4275
Db	4231	TGTCTGATACCGGTGCACCTTTGGAACCGCAGGCATCCAAATGTCTGAAATGCAGGTATG	4290
QY	4276	TGTCAACCCCAAGTGTCGACGATGCTGACGCCATGCGCGCTTGCTGCTGAATATAGCC	4335
Db	4291	TGTCAACCCCAAGTGTCGACGATGCTGACGCCATGCGCGCTTGCTGCTGAATATAGCC	4350
QY	4336	ACAACATTCACCGAGGCTTCTGCGCTGACCAAAATGAACTCCCAAGTCTCCAGACCAAG	4395
Db	4351	ACAACATTCACCGAGGCTTCTGCGCTGACCAAAATGAACTCCCAAGTCTCCAGACCAAG	4410
QY	4396	GAGCCGACGACAGCTTGACCTTGAAAGGCTGATGAAAGTCCGACAGAAATTAACAAACGA	4455
Db	4411	GAGCCGACGACAGCTTGACCTTGAAAGGCTGATGAAAGTCCGACAGAAATTAACAAACGA	4470
QY	4456	GGAACAGCAAGGCTGGGACAGAAATTAATGTTCTCTGAGAGGATTCAAAAGTCTCATTTAT	4515
Db	4471	GGAACAGCAAGGCTGGGACAGAAATTAATGTTCTCTGAGAGGATTCAAAAGTCTCATTTAT	4530
QY	4516	GACAAATGAAGCAAGAACTGACACAGAGCCGCTGAAAGAAATTTGAGCTGTACCTTCCC	4575
Db	4531	GACAAATGAAGCAAGAACTGACACAGAGCCGCTGAAAGAAATTTGAGCTGTACCTTCCC	4590
QY	4576	GACGGGAGTGTATCTATTCAATGTGCGGTTGGTGCTTCCGAATCTCGCAATACAGCCAA	4635

Db	4591	GAACGGGGAATGTATATTCATGATGGCCGTTGGTCTTCGGAATCGCAATATACAGCAAA	465
Qy	4636	GGAGATGTCCCATACATACATGAAGATGAATCTCAACCCGCAACCAACCTGCGCCCGGG	4691
Db	4651	GCA-----	465
Qy	4696	AGAACCCCTTACTGTGCTAGCTCCAGCTTCCCTGACAAACAGCGCTGGGTCAACGCTTA	4751
Db	4654	-----	465
Qy	4756	GAATCAGTTTGGCAGAGTGGAGAGTTTCTAGGGAAAAACGAGAGCTGATCTTAACTG	4811
Db	4654	-----GAAAAAGCGAAGCTGATGCTTAACCTG	468
Qy	4816	CTTGGAAAATCCCTGCTGAAACTGGAAAGGTATGACCGCTTATGACATGAACTGACGCGCT	4875
Db	4681	CTTGGAAATCTCCCTGCTGAAAATGGAAAGGTATGACCGCTTATGACATGAACTGACGCGCT	4744
Qy	4876	CCCTTCAGTGAACAGGTGGTGTGTGGTGGGACCGAGAAAGGGCTTACGCGCTGATGTGC	4933
Db	4741	CCCTTCAGTGAACAGGTGGTGTGTGGTGGGACCGAGAAAGGGCTTACGCGCTGATGTGC	4800
Qy	4936	TTGAAAAATCCCTTAAACCATATGTCCAGAAATTGGAGCATCTTCCAAATTTATATATC	4995
Db	4801	TTGAAAAATCCCTTAAACCATATGTCCAGAAATTGGAGCATCTTCCAAATTTATATATC	4860
Qy	4996	AAGGACCTGGAGAACTCTACTATGATACAGAGAAAGACGGGACCTGTGCTTGTGACAC	5055
Db	4861	AAGGACCTGGAGAACTCTACTATGATACAGAGAAAGACGGGACCTGTGCTTGTGACAC	4920
Qy	5056	GTGAAGAAAGTGAACAGTCCCTGGCCAGTCCCACTGCTGCCACCCGACATCTCA	5115
Db	4921	GTGAAGAAAGTGAACAGTCCCTGGCCAGTCCCACTGCTGCCACCCGACATCTCA	4980
Qy	5116	CCCAACATTTTGAAGCTGTCAAGGGCTGCCACTTTGTTGGGGCAGGCAAGATTGAGAAC	5175
Db	4981	CCCAACATTTTGAAGCTGTCAAGGGCTGCCACTTTGTTGGGGCAGGCAAGATTGAGAAC	5040
Qy	5176	GGGCTCTGCACTGTGAGAGCCATGCGCAAGAGTCCGCAATTCCTCCGTACAAAGAAAC	5235
Db	5041	GGGCTCTGCACTGTGAGAGCCATGCGCAAGAGTCCGCAATTCCTCCGTACAAAGAAAC	5100
Qy	5236	CTCAGCAATATCTGCATCCGGAAAGAGATAGACACTTACAGAGCCCTGACGCTGATCCAC	5295
Db	5101	CTCAGCAATATCTGCATCCGGAAAGAGATAGACACTTACAGAGCCCTGACGCTGATCCAC	5160
Qy	5296	TTCAACCAATTAACGATATCTCTATTTGAAACCAATTAATTTCTACGAATTCGACATGAACAG	5355
Db	5161	TTCAACCAATTAACGATATCTCTATTTGAAACCAATTAATTTCTACGAATTCGACATGAACAG	5220
Qy	5356	TACACGCTCGAGGAATTCCTGGAATTAAGATTAACAATTCCTTGGGACCTGCTGTGTTGGCC	5415
Db	5221	TACACGCTCGAGGAATTCCTGGAATTAAGATTAACAATTCCTTGGGACCTGCTGTGTTGGCC	5280
Qy	5416	GCCTCTTCCAAACAAGCTTCCCTGTCTCAATCGTGCAGGTGAACAACGACGGGCAACGAGAG	5475
Db	5281	GCCTCTTCCAAACAAGCTTCCCTGTCTCAATCGTGCAGGTGAACAACGACGGGCAACGAGAG	5340
Qy	5476	GAGTACTTGTCTGTATTTCCACGAATTTGAGATGTTGCTGATTTCTTACGAAAGCGTAGC	5535
Db	5341	GAGTACTTGTCTGTATTTCCACGAATTTGAGATGTTGCTGATTTCTTACGAAAGCGTAGC	5400
Qy	5536	CGACACAGCAATCTCAAGTGGAGTCCGCTTACCTTTGGCCTTTGGCTTACAGAAACCTTAT	5595
Db	5401	CGACACAGCAATCTCAAGTGGAGTCCGCTTACCTTTGGCCTTTGGCTTACAGAAACCTTAT	5460
Qy	5596	CTGTTTGTGACCCACTTCAACTCACTCACTGGAAGTAAATTAGATTCAGGACCGCTCTCAGCA	5655
Db	5461	CTGTTTGTGACCCACTTCAACTCACTCACTGGAAGTAAATTAGATTCAGGACCGCTCTCAGCA	5520
Qy	5656	GGAGACCCCTGCGAGCGTACTTGGACATCCGGAACCGCGCTTACTTGGGCCCTTGCCATT	5715

Db 5521 GGGACCCCTCCGAGGCTACTGACATCCCGAACCCGCGCTACTGGGCCCTGCCATT 5580
Qy 5716 TCCTCAGAGAGATTTACTTGGCTCTCATACAGAGATAATTAAGGATCATTTGCTGC 5775
Db 5581 TCCTCAGAGAGATTTACTTGGCTCTCATACAGAGATAATTAAGGATCATTTGCTGC 5640
Qy 5776 AAGGAAACTCTGTGAAGAGTCCGGCACTGAACAACAACCGGGGCCCTGCCACTCCCGC 5835
Db 5641 AAGGAAACTCTGTGAAGAGTCCGGCACTGAACAACAACCGGGGCCCTGCCACTCCCGC 5700
Qy 5836 AGCAGCCCAACAAGAGAGGCGCCACGATCAACAGACAATCACCAAGCGGTGGCC 5895
Db 5701 AGCAGCCCAACAAGAGAGGCGCCACGATCAACAGACAATCACCAAGCGGTGGCC 5760
Qy 5896 TCCAGCCCAAGCGCCCGCCGAGAGCCCAAGCCCGGAGAGAGCAACCCACCGC 5955
Db 5761 TCCAGCCCAAGCGCCCGCCGAGAGCCCAAGCCCGGAGAGCAACCCACCGC 5820
Qy 5956 TACCGCAGAGGGGCGAGCCGAGCTGCGAGAGGACAAGTCTCTGGCCCGCCCTGAGCGA 6015
Db 5821 TACCGCAGAGGGGCGAGCCGAGCTGCGAGAGGACAAGTCTCTGGCCCGCCCTGAGCGA 5880
Qy 6016 GAGAAGTCTCCCGCCGAGTACTCAGACACGGGAGAGAGCGGTCTCCCGGAGCGTCTTT 6075
Db 5881 GAGAAGTCTCCCGCCGAGTACTCAGACACGGGAGAGAGCGGTCTCCCGGAGCGTCTTT 5940
Qy 6076 GAAGACAGCAGCAGAGGGCGGCTGCTGCGAGAGCCGCTGAGAGACCCGCTGTCCAGGTG 6135
Db 5941 GAAGACAGCAGCAGAGGGCGGCTGCTGCGAGAGCCGCTGAGAGACCCGCTGTCCAGGTG 6000
Qy 6136 AACAGGCTCTGGAGCAAGTCTTC 6158
Db 6001 AACAGGAGAGAGGAGCAAGTGC 6023

RESULT 15
US-10-757-262-51
; Sequence 51, Application US/10757262
; Publication No. US20040197825A1
; GENERAL INFORMATION:
; APPLICANT: Karicheci, Venkateswarlu
; APPLICANT: Slioh-Santlago, Inmaculada
; APPLICANT: Eliasof, Scott D.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING
; TITLE OF INVENTION: UROLOGICAL DISORDERS USING 44390, 54181, 211, 5687, 884,
; TITLE OF INVENTION: 1405, 636, 4421, 5410, 30905, 2045, 16405, 18500, 2047,
; TITLE OF INVENTION: 33751, 52872, 14063, 20739, 32544, 43239, 44373, 51164,
; TITLE OF INVENTION: 53010, 16852, 1587, 2207, 22245, 2387, 52908, 69112, 14990,
; TITLE OF INVENTION: 18547, 115, 579, 15985, 15625, 760, 18603, 2395, 2554, 8675,
; TITLE OF INVENTION: 32720, 4809, 14303, 16816, 17827, 32620, 577, 619, 1423,
; TITLE OF INVENTION: 2158, 8263, 15402, 16209, 16386, 21165, 30911, 41897, 1643,
; TITLE OF INVENTION: 2543, 9626, 13231, 32409, 84260, 2882, 8203, 32678 OR
; TITLE OF INVENTION: 55053
; FILE REFERENCE: MP103-007P1RNONMIN
; CURRENT APPLICATION NUMBER: US/10757, 262
; CURRENT FILING DATE: 2004-01-14
; PRIOR APPLICATION NUMBER: US 60/440, 318
; PRIOR FILING DATE: 2003-01-15
; PRIOR APPLICATION NUMBER: US 60/444, 783
; PRIOR FILING DATE: 2003-02-04
; PRIOR APPLICATION NUMBER: US 60/457, 901
; PRIOR FILING DATE: 2003-03-27
; PRIOR APPLICATION NUMBER: US 60/468, 775
; PRIOR FILING DATE: 2003-05-08
; PRIOR APPLICATION NUMBER: US 60/471, 614
; PRIOR FILING DATE: 2003-05-19
; PRIOR APPLICATION NUMBER: US 60/478, 742
; PRIOR FILING DATE: 2003-06-16
; PRIOR APPLICATION NUMBER: US 60/488, 529
; PRIOR FILING DATE: 2003-07-18
; PRIOR APPLICATION NUMBER: US 60/491, 156
; PRIOR FILING DATE: 2003-07-30
; PRIOR APPLICATION NUMBER: US 60/499, 594

; PRIOR FILING DATE: 2003-09-02
; PRIOR APPLICATION NUMBER: US 60/506, 332
; PRIOR FILING DATE: 2003-09-26
; NUMBER OF SEQ ID NOS: 136
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 51
; LENGTH: 6574
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (19)...(6180)
US-10-757-262-51

Query Match 91.8%; Score 5661.4; DB 18; Length 6574;
Best Local Similarity 95.8%; Pred. No. 0;
Matches 5944; Conservative 0; Mismatches 16; Indels 243; Gaps 3;

Qy 1 ATGTTGAAGTTCAATATGAGAGCGGGAATCTTTGATGCTGTGCTGTAACCAATT 60
Db 19 ATGTTGAAGTTCAATATGAGAGCGGGAATCTTTGATGCTGTGCTGTAACCAATT 78
Qy 61 GCCAGCCGGGCTCCAGGCTGAATCTGTTCTCCAGGGGAAACACCTTTAGACTCA 120
Db 79 GCCAGCCGGGCTCCAGGCTGAATCTGTTCTCCAGGGGAAACACCTTTAGACTCA 138
Qy 121 CAGCAGATGCTCTCTCTTCCCGAAGAGATTAAGATCCCTCTTGTCTCTTGA 180
Db 139 CAGCAGATGCTCTCTCTTCCCGAAGAGATTAAGATCCCTCTTGTCTCTTGA 198
Qy 181 GAATGAGTACGCTCTCTGATGAAGATTAAAGACGTGAGCACTTTGTCGGAATAT 240
Db 199 GAATGAGTACGCTCTCTGATGAAGATTAAAGACGTGAGCACTTTGTCGGAATAT 258
Qy 241 TCCGACACCATAGCTGATTAAGAGAGCTCAAGCTTCGGCAAGAGCTTGAAGTGA 300
Db 259 TCCGACACCATAGCTGATTAAGAGAGCTCAAGCTTCGGCAAGAGCTTGAAGTGA 318
Qy 301 AGCTTTAGATTTGTGCTCACTTGTGTAAGTGAAGTGAAGTGAAGTGAAGTGA 360
Db 319 AGCTTTAGATTTGTGCTCACTTGTGTAAGTGAAGTGAAGTGAAGTGAAGTGA 378
Qy 361 GACATCTATGCTATGAAGATGATGAAGAGAGCTTTATTTGCTCCAGAGAGCTTCA 420
Db 379 GACATCTATGCTATGAAGATGATGAAGAGAGCTTTATTTGCTCCAGAGAGCTTCA 438
Qy 421 TTTTGTGAGGAAGCGGAACTATTTATCTGAGAGCAAGCCCGTGGATCCCAATTA 480
Db 439 TTTTGTGAGGAAGCGGAACTATTTATCTGAGAGCAAGCCCGTGGATCCCAATTA 498
Qy 481 CAGTATGCTTTGAGAGCAAAATCACCTTATCTGATGAAGTGAAGTGAAGTGA 540
Db 499 CAGTATGCTTTGAGAGCAAAATCACCTTATCTGATGAAGTGAAGTGAAGTGA 558
Qy 541 GACTTGCTGCTCACTTTTGAATGATGAGAGCAAGTTAAGTGAAGTGAAGTGA 600
Db 559 GACTTGCTGCTCACTTTTGAATGATGAGAGCAAGTTAAGTGAAGTGAAGTGA 618
Qy 601 TACTTACTGAGCTATTTTGGCTGTTCAAGAGCTTATCTGATGAGTATGATG 660
Db 619 TACTTACTGAGCTATTTTGGCTGTTCAAGAGCTTATCTGATGAGTATGATG 678
Qy 661 GACATCAAGGCTGAGACATTTCTGTTGACCGACAGACATCAAGCTGTGATTTT 720
Db 679 GACATCAAGGCTGAGACATTTCTGTTGACCGACAGACATCAAGCTGTGATTTT 738
Qy 721 GGAATCTGCGGAAATGAATTCAAACAAGATGTAATGCCAACTCCGATTTGGAGC 780
Db 739 GGAATCTGCGGAAATGAATTCAAACAAGATGTAATGCCAACTCCGATTTGGAGC 798
Qy 781 CCAGATTACATGCTCTCTGAAGTCTGATGTAAGCGGAGTGAAGAAAGGACCTTAC 840
Db 799 CCAGATTACATGCTCTCTGAAGTCTGATGTAAGCGGAGTGAAGAAAGGACCTTAC 858

QY 841 GGCCTGAGCTGTGATCGTGTGCTCAGTGGGCGTGATGTGCTTATGAGTATTTATGGAGA 900
DB 859 GGCCTGAGCTGTGATCGTGTGCTCAGTGGGCGTGATGTGCTTATGAGTATTTATGGAGA 918
QY 901 TCCCTTCGAGAGGGGAACTCTGCGAGAACTTTCATTAACATTTATGATTTTCCAGCGG 960
DB 919 TCCCTTCGAGAGGGGAACTCTGCGAGAACTTTCATTAACATTTATGATTTTCCAGCGG 978
QY 961 TTTTGAATTTCCAGATGACCCCAAGTGAAGTGAAGTCTTCTGTAATCTGATTCAGAGC 1020
DB 979 TTTTGAATTTCCAGATGACCCCAAGTGAAGTGAAGTCTTCTGTAATCTGATTCAGAGC 1038
QY 1021 TTGTGTGGCGGCGAGAAAGAGAGCTGAAGTTTGAAGTCTTGTCTGCTCCTTTCTTC 1080
DB 1039 TTGTGTGGCGGCGAGAAAGAGAGCTGAAGTTTGAAGTCTTGTCTGCTCCTTTCTTC 1098
QY 1081 TCTTAAATTTGATGGAACAATTCGTAACTCTCTCCCTCCCTGTTTCCACCTCAG 1140
DB 1099 TCTTAAATTTGATGGAACAATTCGTAACTCTCTCCCTCCCTGTTTCCACCTCAG 1158
QY 1141 TCTGACGATGACCTTCAATTTTGAATGAACAGAGAAATTCGTGGTTTCACTCTCT 1200
DB 1159 TCTGACGATGACCTTCAATTTTGAATGAACAGAGAAATTCGTGGTTTCACTCTCT 1218
QY 1201 CCGTGCAGCTGAGCCCTCAGGCTTCTCGGCTGAAGAACTGCCCTTTGTGGGTTTTCG 1260
DB 1219 CCGTGCAGCTGAGCCCTCAGGCTTCTCGGCTGAAGAACTGCCCTTTGTGGGTTTTCG 1278
QY 1261 TACAGCAAGGCACTGGGGAATTTCTGTGATGATCTGAGTCTGTGTGTGGGCTGGAATCC 1320
DB 1279 TACAGCAAGGCACTGGGGAATTTCTGTGATGATCTGAGTCTGTGTGTGGGCTGGAATCC 1338
QY 1321 CTTGCCAAGACTAGCTCCATGAGAAAGAACTTCTCATCAAAAGCAAGAGCTACAGAC 1380
DB 1339 CTTGCCAAGACTAGCTCCATGAGAAAGAACTTCTCATCAAAAGCAAGAGCTACAGAC 1398
QY 1381 TCTCAGAGCAAGTGTCAAGATGAGAGCAAGAAATGACCCGTTTCACTCGAGAGTGTCA 1440
DB 1399 TCTCAGAGCAAGTGTCAAGATGAGAGCAAGAAATGACCCGTTTCACTCGAGAGTGTCA 1458
QY 1441 GAGGTGAGGCTGTGCTTGTAGTCAAGAGAGAGTGAAGGCTCTGAGACTCAGAGA 1500
DB 1459 GAGGTGAGGCTGTGCTTGTAGTCAAGAGAGAGTGAAGGCTCTGAGACTCAGAGA 1518
QY 1501 TCCCTCTTGAGAGAGAACTTGTCTACCTTACATCAAGATGAGTCTTAAAGCGAAGT 1560
DB 1519 TCCCTCTTGAGAGAGAACTTGTCTACCTTACATCAAGATGAGTCTTAAAGCGAAGT 1578
QY 1561 TTGGAGCAAGCAAGGATGAGAGTGTCTCCAGAGAGATGACAAAGCACTGAGCTTCTCAT 1620
DB 1579 TTGGAGCAAGCAAGGATGAGAGTGTCTCCAGAGAGATGACAAAGCACTGAGCTTCTCAT 1638
QY 1621 GATATCAGAGAGAGAGCGGAGAACTCAAGAAATCAAGAGAGAGATCAAGAGCTTCAA 1680
DB 1639 GATATCAGAGAGAGAGCGGAGAACTCAAGAAATCAAGAGAGAGATCAAGAGCTTCAA 1698
QY 1681 GTGGAAGAAATGAGTGTGATGATCAAGTGAAGAGATCTTGTCTCAGCAAGAGA 1740
DB 1699 GTGGAAGAAATGAGTGTGATGATCAAGTGAAGAGATCTTGTCTCAGCAAGAGA 1758
QY 1741 CCGAGTGTATCTTACGAATCTGAGCTGAGAGTCTCGGCTTGTCTGAGAGATTCAG 1800
DB 1759 CCGAGTGTATCTTACGAATCTGAGCTGAGAGTCTCGGCTTGTCTGAGAGATTCAG 1818
QY 1801 CCGAAGAGCAAGATGTGAGATTAACGTTGAAGGCTAAGATCAAGAGAGCTTGA 1860
DB 1819 CCGAAGAGCAAGATGTGAGATTAACGTTGAAGGCTAAGAGATCAAGAGAGCTTGA 1878
QY 1861 GTGGAGAGATATGCAAACTGAGAGAGATCAATGTCTGAGCAGAGCTCAAAATTCAGAG 1920
DB 1879 GTGGAGAGATATGCAAACTGAGAGAGATCAATGTCTGAGCAGAGCTCAAAATTCAGAG 1938
QY 1921 CTCGAAGAGAACTGAGAGAGCTGTATTAAGCCAGACGAGAGCTCAGCTGCTCAG 1980
DB 1939 CTCGAAGAGAACTGAGAGAGCTGTATTAAGCCAGACGAGAGCTCAGCTGCTCAG 1962
QY 1981 AATATCCGCAAGGCAAG 2040
DB 1963 -----GCAAG 2010
QY 2041 GATTCCTTCTGAAGGATCAG 2100
DB 2011 GATTCCTTCTGAAGGATCAG 2070
QY 2101 GAGAACAGGATTAAGAGACTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2160
DB 2071 GAGAACAGGATTAAGAGACTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2130
QY 2161 ATCCAGCAAAATCCCAACAGATCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2220
DB 2131 ATCCAGCAAAATCCCAACAGATCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2190
QY 2221 AAACATCGGAGAGCCCAAGTCTCAGCCAGACCTAGAAAGTGCACCTGAAGAGAGAGAG 2280
DB 2191 AAACATCGGAGAGCCCAAGTCTCAGCCAGACCTAGAAAGTGCACCTGAAGAGAGAGAG 2250
QY 2281 CAGCACTATGAGAGAAAGATTAAGTGTGAGCAATCAATTAAGAGAGAGAGAGAGAGAG 2340
DB 2251 CAGCACTATGAGAGAAAGATTAAGTGTGAGCAATCAATTAAGAGAGAGAGAGAGAGAG 2310
QY 2341 AAGGAGACCTGAGAGAAATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2400
DB 2311 AAGGAGACCTGAGAGAAATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2370
QY 2401 ATTCTAGGGAACAG 2460
DB 2371 ATTCTAGGGAACAG 2430
QY 2461 CAGAGATTTGTGAGACTGTCTGAGAGCAATTAAGTGTGAGCAATTAAGAGAGAGAGAG 2520
DB 2431 CAGAGATTTGTGAGACTGTCTGAGAGCAATTAAGTGTGAGCAATTAAGAGAGAGAGAG 2490
QY 2521 CAAAGGAGATGAG 2580
DB 2491 CAAAGGAGATGAG 2550
QY 2581 CTGAGAGACAG 2640
DB 2551 CTGAGAGACAG 2610
QY 2641 AAGATCAGCCCAAGAGACCAAGTGAAGAAATCGGCTGTGGAATCTGAGAGCAAGATTTG 2700
DB 2611 AAGATCAGCCCAAGAGACCAAGTGAAGAAATCGGCTGTGGAATCTGAGAGCAAGATTTG 2670
QY 2701 CCGGAGGTCAGTCTGAG 2760
DB 2671 CCGGAGGTCAGTCTGAG 2730
QY 2761 CTACAGCTCTCCCTCAG 2820
DB 2731 CTACAGCTCTCCCTCAG 2790
QY 2821 GCCCTGAGAGAGCACTTTCGCAAGGCGAAGCAGAGCTGGAAGAGACCAAGCAGAGAGCT 2880
DB 2791 GCCCTGAGAGAGCACTTTCGCAAGGCGAAGCAGAGCTGGAAGAGACCAAGCAGAGAGCT 2850
QY 2881 GAGAGAGAGATCCAGAGCACTCAGGCAATGAGATGAAGTCAAGCGCAAAATTTGATCTCT 2940
DB 2851 GAGAGAGAGATCCAGAGCACTCAGGCAATGAGATGAAGTCAAGCGCAAAATTTGATCTCT 2910
QY 2941 CTTCTGTAACAGCTGTATCTGTAATCAAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3000
DB 2911 CTTCTGTAACAGCTGTATCTGTAATCAAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2970
QY 3001 GACAAAGCTGAATCAACCAAACTTCTATTGTCAAACAACTGATGAGAGCTTCT 3060

Db 2971 GACAAAGCTAACTCAACAAACAAACTTCTACTTGTCCAAACAATCGATGAGCTTCT 3030
Qy 3061 GGGCCCAACGACGAGATTTGTCAAACTGCAAGTGAAGTCCATCTCCGCGGAGATC 3120
Db 3031 GGGGCCAACGACGAGATTTGTCAAACTGCAAGTGAAGTCCATCTCCGCGGAGATC 3090
Qy 3121 ACCGAAACGAGAGATGACGCTTACCAAGCAGAAACGATGAGGCTTGAAAGCAACG 3180
Db 3091 ACCGAAACGAGAGATGACGCTTACCAAGCAGAAACGATGAGGCTTGAAAGCAACG 3150
Qy 3181 TGCACCATGCTGAGAGAAACAGTCAATGATTTTGAAGGCTTAAACGATGAGCTGTGAA 3240
Db 3151 TGCACCATGCTGAGAGAAACAGTCAATGATTTTGAAGGCTTAAACGATGAGCTGTGAA 3210
Qy 3241 AAAGAGGCGAGTGGAGAGCCTGAGAGAGGCTCTGGGTGATGAGAAATCCAGTTTGA 3300
Db 3211 AAAGAGGCGAGTGGAGAGCCTGAGAGAGGCTCTGGGTGATGAGAAATCCAGTTTGA 3270
Qy 3301 TGTGCGGTTGAGAGCTGCAAGAGATCTGGAACCCAGAAACAGACGAGCGGAGAGCC 3360
Db 3271 TGTGCGGTTGAGAGCTGCAAGAGATCTGGAACCCAGAAACAGACGAGCGGAGAGCC 3330
Qy 3361 GATCAGCGGATCAACCGAGTCTCCGACAGTGTGAGAGCTGCGAGTGAAGAGCAAGGCT 3420
Db 3331 GATCAGCGGATCAACCGAGTCTCCGACAGTGTGAGAGCTGCGAGTGAAGAGCAAGGCT 3390
Qy 3421 GAGATTCTGCTGCGACGAGGCTCTCAAAAGCAGAAAGCTGAAGGCGAGAGCTCTCT 3480
Db 3391 GAGATTCTGCTGCGACGAGGCTCTCAAAAGCAGAAAGCTGAAGGCGAGAGCTCTCT 3450
Qy 3481 GACAAAGCTCAATGACCTGGAAGAAACATGCTATGCTTGAATGAATGCGCGAAGCTTA 3540
Db 3451 GACAAAGCTCAATGACCTGGAAGAAACATGCTATGCTTGAATGAATGCGCGAAGCTTA 3510
Qy 3541 CAGCAGAACTGGAAGCTGAAACGAGAGCTCAACAGAGGCTTCTGGAAGAGCAAGCCAA 3600
Db 3511 CAGCAGAACTGGAAGCTGAAACGAGAGCTCAACAGAGGCTTCTGGAAGAGCAAGCCAA 3570
Qy 3601 TTAACAGAGAGATGACCTGCAAGAAATCACTTTCCGTCGATCAAGGACCTGCA 3660
Db 3571 TTAACAGAGAGATGACCTGCAAGAAATCACTTTCCGTCGATCAAGGACCTGCA 3630
Qy 3661 GAAGCTCTAGATCGGCTGATCTACTGAAGACAGAAAGAGTGAAGTCAAGTCAAGCTG 3720
Db 3631 GAAGCTCTAGATCGGCTGATCTACTGAAGACAGAAAGAGTGAAGTCAAGTCAAGCTG 3690
Qy 3721 GAAAACATTCAAGTTCTCTATTTCTCATGAAAAGGTGAAAATGGAAGCACTATTTCTCA 3780
Db 3691 GAAAACATTCAAGTTCTCTATTTCTCATGAAAAGGTGAAAATGGAAGCACTATTTCTCA 3750
Qy 3781 CAAACCAAACTCATGATTTTCTGCAAGCCAAAATGGAACCACTGCTTAAAAAAGAAA 3838
Db 3751 CAAACCAAACTCATGATTTTCTGCAAGCCAAAATGGAACCACTGCTTAAAAAAGAAA 3810
Qy 3839 -----AGTTTCTCTGCACTAC 3855
Db 3811 GGTTTATTTAGTGAACGAAAAGAGACCTGCTTACCAACAGGTTCTCTGCACTAC 3870
Qy 3856 AATGAGCTGAGCTGAGCTGAGAAAGAGAAAGCTGCTGTGTGCAAGCTGAGAAAGCC 3915
Db 3871 AATGAGCTGAGCTGAGCTGAGAAAGAGAAAGCTGCTGTGTGCAAGCTGAGAAAGCC 3930
Qy 3916 CTTTCAGAGAACCGGACATCGAGCTCCGCGCCGCGAGAGAAAGTGGCCACCCGAAAGCA 3975
Db 3931 CTTTCAGAGAACCGGACATCGAGCTCCGCGCGCGAGAGAAAGTGGCCACCCGAAAGCA 3990
Qy 3976 ACCGACCAACCAACCACTCAACGAGCCAGCCAGCCAGAGGAGAGATGCGCATGTCGCC 4035
Db 3991 ACCGACCAACCAACCACTCAACGAGCCAGCCAGCCAGAGGAGAGATGCGCATGTCGCC 4050
Qy 4036 ATGCTGCGGTGCGCAAGAGCAACGACCAAGTGCATGAGCTGCTGCGCCCGCATCAAGC 4095

Db 4051 ATGTCGCTGCGCAGAGACCAAGCCAGTGCATGAGCTGTGCGCCCGCATCCAGC 4110
Qy 4096 CGCAGAAAGAGTCTTCAACTCCAGAGAAATTTAGTGGCGTCTTAAAGAACCGATCAC 4155
Db 4111 CGCAGAAAGAGTCTTCAACTCCAGAGAAATTTAGTGGCGTCTTAAAGAACCGATCAC 4170
Qy 4156 CACAAATTTCTCAACGATTTCAACGTAAGCTGAACATGGAACCCAAAGTGTGTG 4215
Db 4171 CACAAATTTCTCAACGATTTCAACGTAAGCTGAACATGGAACCCAAAGTGTGTG 4230
Qy 4216 TGTCTGGAATCCGTGACCTTTGAGACGCAAGGATCCAAATGTCTGAAATGTCAGTATG 4275
Db 4231 TGTCTGGAATCCGTGACCTTTGAGACGCAAGGATCCAAATGTCTGAAATGTCAGTATG 4290
Qy 4276 TGTCAACCCCAAGTCTCAACGTCCTTGCAAGCCACCTGCGGCTTGCCTGTAATATGCC 4335
Db 4291 TGTCAACCCCAAGTCTCAACGTCCTTGCAAGCCACCTGCGGCTTGCCTGTAATATGCC 4350
Qy 4336 ACAACCTTCAACGAGGCTTCTGCGTGAACAAATGAATCTCCCAAGTCTCCAGACCAAG 4395
Db 4351 ACAACCTTCAACGAGGCTTCTGCGTGAACAAATGAATCTCCCAAGTCTCCAGACCAAG 4410
Qy 4396 GAGCCGACGAGCCTTGAACCTGGAAGGCTGATGAGGCTGCCAGAAATTAACAAACGA 4455
Db 4411 GAGCCGACGAGCCTTGAACCTGGAAGGCTGATGAGGCTGCCAGAAATTAACAAACGA 4470
Qy 4456 GGAACAGAGGCTGGGACAGAGATGACATTTGCTGAGAGGATCAAAAGTCTCATTTAT 4515
Db 4471 GGAACAGAGGCTGGGACAGAGATGACATTTGCTGAGAGGATCAAAAGTCTCATTTAT 4530
Qy 4516 GACAAATGAAGCCAGAGAGCTGGAACAGAGGCGGCTGGAAGAAATTTGAGCTGTGCTTCC 4575
Db 4531 GACAAATGAAGCCAGAGAGCTGGAACAGAGGCGGCTGGAAGAAATTTGAGCTGTGCTTCC 4590
Qy 4576 GAGGGGAGTATCTATTAATCAATGAGCGTGTGAGCTTCCGAATCGGAAATTAACAGCCAA 4635
Db 4591 GAGGGGAGTATCTATTAATCAATGAGCGTGTGAGCTTCCGAATCGGAAATTAACAGCCAA 4650
Qy 4636 GCAGATGTCCCATATCACTAGAAATGSAATCTCACCCGACACCACTGCTGGCCCGGG 4695
Db 4651 GCA----- 4653
Qy 4696 AGAACCTCTTACTTGTCTAGCTCCAGCTTCTGCAAAACAGCGCTGGTCAACCGCTTA 4755
Db 4654 ----- 4653
Qy 4756 GAATCAATTGTCCGAGGTGGAGAGTTTCTAGGAAAAGACAGAGCTGATGCTAAACCTG 4815
Db 4654 -----GAAAAGCAGAAAGCTGATGCTAAACCTG 4680
Qy 4816 CTTGAAAACCTCCGCTGAACCTGAAAGTGAAGTGAACCGTCTAGACATGAACCTGACGCTG 4875
Db 4681 CTTGAAAACCTCCGCTGAACCTGAAAGTGAAGTGAACCGTCTAGACATGAACCTGACGCTG 4740
Qy 4876 CCCTTCAGTACCAAGTGTGTGTGTGGGACCGAGAGAGGCTCTACGCTGAAATGTC 4935
Db 4741 CCCTTCAGTACCAAGTGTGTGTGTGGGACCGAGAGAGGCTCTACGCTGAAATGTC 4800
Qy 4936 TTGAAAACCTCCCTAACCTCAATGTCCTCAGAGAAATTTGAGCAATCTTCCAAATTAATATC 4995
Db 4801 TTGAAAACCTCCCTAACCTCAATGTCCTCAGAGAAATTTGAGCAATCTTCCAAATTAATATC 4860
Qy 4996 AAGGACCTGGAAGAGTCACTCATGATGACAGAGAAAGGAGGCACTGTGTCTTGTGAGC 5055
Db 4861 AAGGACCTGGAAGAGTCACTCATGATGACAGAGAAAGGAGGCACTGTGTCTTGTGAGC 4920
Qy 5056 GTGAGAAAGTGAAGAGTCCCTGAGCCAGTCCCACTGCTGCGCAGCCGACATCTCA 5115
Db 4921 GTGAGAAAGTGAAGAGTCCCTGAGCCAGTCCCACTGCTGCGCAGCCGACATCTCA 4980
Qy 5116 CCCAACTTTTGAAGTGTCAAGGCTGCACTTGTGTTGGGCAAGGCAAGATTTGANAAC 5175
Db 4981 CCCAACTTTTGAAGTGTCAAGGCTGCACTTGTGTTGGGCAAGGCAAGATTTGANAAC 5040

QY 5176 GGGCTTCGATCTGTGCGAGCCATGCGCAAGTCGTCAATTCTCCGCTACACGAAAC 5235
DB 5041 GGGCTTCGCACTCTGTGAGCCATGCCAGCAAGTCGTCAATTCTCCGCTACACGAAAC 5100
QY 5236 CTGAGAAATCTGCAATCCGGAAGAGTGAACCTCAGAGCCCTGAGCTGATCCAC 5295
DB 5101 CTGAGCAAAATCTGCAATCCGGAAGAGTGAACCTCAGAGCCCTGAGCTGATCCAC 5160
QY 5296 TTGACCAATTACAGTATCTCATTTGGAACCAATTAATTTCTACGAATGACATGAACAG 5355
DB 5161 TTGACCAATTACAGTATCTCATTTGGAACCAATTAATTTCTACGAATGACATGAACAG 5220
QY 5356 TACACGCTCGAGGAATCTGAGATGAATGACCAATCTTGCGACCTGCTGTGTTGCC 5415
DB 5221 TACACGCTCGAGGAATCTGAGATGAATGACCAATCTTGCGACCTGCTGTGTTGCC 5280
QY 5416 GCGCTCTTCCAAAGCTTCCCTGTCTCAATGTGCAAGTGAACAGGCGAGCGGAGAG 5475
DB 5281 GCGCTCTTCCAAAGCTTCCCTGTCTCAATGTGCAAGTGAACAGGCGAGCGGAGAG 5340
QY 5476 GAGTACTGTGCTGTGTTCCAGAAATTGAGTGTTCGTGATTTCTTACGGAAGACTAGC 5535
DB 5341 GAGTACTGTGCTGTGTTCCAGAAATTGAGTGTTCGTGATTTCTTACGGAAGACTAGC 5400
QY 5536 CGCAGACGATCTCAAGTGAAGTCGCTTACCTTTGCGCTTTCCTTACAGAGAACCTAT 5595
DB 5401 CGCAGACGATCTCAAGTGAAGTCGCTTACCTTTGCGCTTTCCTTACAGAGAACCTAT 5460
QY 5596 CTGTTTGTGACCCACTTCACTCACTGSAATTAATGAGATCCAGGACGCTCTCTCAGCA 5655
DB 5461 CTGTTTGTGACCCACTTCACTCACTGSAATTAATGAGATCCAGGACGCTCTCTCAGCA 5520
QY 5656 GGGAGCCCTGCGCGAGGCTGACCTGACATCCGAAACCGCGCTACCTGAGCCCTGCATT 5715
DB 5521 GGGAGCCCTGCGCGAGGCTGACCTGACATCCGAAACCGCGCTACCTGAGCCCTGCATT 5580
QY 5716 TCCTCAGGAGCGATTTACTTGGCGTCTCTATACAGAGATTAATTAAGGTCATTTGCTGC 5775
DB 5581 TCCTCAGGAGCGATTTACTTGGCGTCTCTATACAGAGATTAATTAAGGTCATTTGCTGC 5640
QY 5776 AAGGGAACCTGTAAGAGTCCGGCACTGAAACCAACCGGGCCCGTCCACTCCCGC 5835
DB 5641 AAGGGAACCTGTAAGAGTCCGGCACTGAAACCAACCGGGCCCGTCCACTCCCGC 5700
QY 5836 AGCAGCCCAAGCAGAGGCCCAACCAAGTACAAAGACATCAACCAAGCGCTGAGCC 5895
DB 5701 AGCAGCCCAAGCAGAGGCCCAACCAAGTACAAAGACATCAACCAAGCGCTGAGCC 5760
QY 5896 TCCAGCCCAAGCAGAGGCCCAACCAAGTACAAAGACATCAACCAAGCGCTGAGCC 5955
DB 5761 TCCAGCCCAAGCAGAGGCCCAACCAAGTACAAAGACATCAACCAAGCGCTGAGCC 5820
QY 5956 TACCGGAGGGGCGGACCGAGCTGCGCAGGGAACAATCTCCGCGCCCGCTGAGCGCA 6015
DB 5821 TACCGGAGGGGCGGACCGAGCTGCGCAGGGAACAATCTCCGCGCCCGCTGAGCGCA 5880
QY 6016 GAGAAATCCCGCGCGGATTACTCAGCAGCGGAGAGAGGCTCCCGCGAGGCTGTTT 6075
DB 5881 GAGAAATCCCGCGCGGATTACTCAGCAGCGGAGAGAGGCTCCCGCGAGGCTGTTT 5940
QY 6076 GAAAGCAGCAGCAGGAGGCGGCTGCTGCGGAGGCGCTGAGGACCCCGCTGTCCAGGTG 6135
DB 5941 GAAAGCAGCAGCAGGAGGCGGCTGCTGCGGAGGCGCTGAGGACCCCGCTGTCCAGGTG 6000
QY 6136 AACAGGTCCTGGAGCAAGTCTTC 6158
DB 6001 AACAGGGAAGGCGCAGAGTGC 6023

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using bw model

Run on: May 16, 2005, 09:05:39 ; Search time 11990 Seconds
(without alignments)
19571.828 Million cell updates/sec

Title: US-10-791-666-1
Perfect score: 6165
Sequence: 1 acgtcgaattcaataatg9.....gggaccagcttcagataa 6165

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues
Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:
1: gb_esc1:*
2: gb_esc2:*
3: gb_hrc:*
4: gb_esc3:*
5: gb_esc4:*
6: gb_esc5:*
7: gb_esc6:*
8: gb_gss1:*
9: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	913.6	14.8	1011	5	BQ070955 AGENCOURT
2	822.6	13.3	879	5	BQ228524 AGENCOURT
3	814	13.2	956	5	BQ071141 AGENCOURT
4	810.2	13.1	830	5	BU181633 AGENCOURT
5	750.6	12.2	849	4	BI253509 602973370
6	732.8	11.9	855	5	BK342268 BX342268
7	732	11.9	757	7	CN281134 170004237
8	705.6	11.4	958	3	BC031156 Mus muscu
9	698.2	11.3	843	4	BC912161 602812833
10	649.4	10.5	653	7	CN295302 170006001
11	629.2	10.2	652	2	BF905370 UI3-MT026
12	628.8	10.2	640	5	BQ807302 NISC_KK01
13	620.4	10.1	716	7	CK636398 UI-M-HNO
14	603.4	9.8	881	4	BC976452 602846269
15	602.6	9.8	698	5	BP146990 BP146990
16	599	9.7	891	4	BI558919 603240369
17	591.4	9.6	746	6	CA749280 UI-M-FV0
18	588.6	9.5	681	2	AM605350 QV3-DT004
19	581.6	9.4	968	5	BU138506 603132657
20	575.2	9.3	1078	5	BU139267 603132236
21	574.4	9.3	698	6	CD348416 UI-M-FV0
22	573.6	9.3	647	7	CF744580 UI-M-GV0
23	570.4	9.2	730	7	CF723360 UI-M-GV0
24	566.4	9.2	1085	5	BM904785 AGENCOURT

25	557.4	9.0	647	2	BB207065	BB207065
26	557.2	9.0	675	7	CF735558	CF735558 UI-M-HB0
27	538	8.7	566	5	BK481181	BK481181 DKF2P686P
28	515.4	8.4	839	7	CN225795	CN225795 WLA078D10
29	511.8	8.3	775	5	BU105833	BU105833 603005490
30	511	8.3	804	5	BU236617	BU236617 603411670
31	508.6	8.2	547	7	CF744826	CF744826 UI-M-GV0
32	498.2	8.1	719	5	BU339218	BU339218 603515216
33	497.4	8.0	563	7	CN536028	CN536028 UI-M-HS0
34	495.6	8.0	933	3	BU232508	BU232508 603408272
35	494.6	8.0	532	7	CR735660	CR735660 CR735660
36	491.6	8.0	875	5	BK723780	BK723780 BX723780
37	489.6	7.9	525	7	CR735680	CR735680 CR735680
38	489.6	7.9	596	7	CF540358	CF540358 UI-M-GV0
39	488	7.9	657	5	BQ831488	BQ831488 IL61n2149
40	481	7.8	820	5	BQ941506	BQ941506 AGENCOURT
41	479.6	7.8	922	7	CR746615	CR746615 CR746615
42	460.8	7.5	798	5	BX888120	BX888120 BX888120
43	457.6	7.4	789	7	CO798747	CO798747 AGENCOURT
44	451.4	7.3	459	5	BU429033	BU429033 UI-HF-BNO
45	450	7.3	600	4	BG808196	BG808196 2082-86 M

ALIGNMENTS

RESULT 1
LOCUS BQ070955 1011 bp mRNA linear EST 02-APR-2002
DEFINITION AGENCOURT_6855647 NIH_MGC_47 Homo sapiens cDNA clone IMAGE:5923441
5', mRNA sequence.
ACCESSION BQ070955
VERSION BQ070955.1 GI:13900001
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE NIH-MGC http://mgi.nci.nih.gov/
1 (bases 1 to 1011)
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE Unpublished (1999)
JOURNAL Contact: Robert Strausberg, Ph.D.
COMMENT Email: cgabbs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LINT)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LINT at:
http://image.llnl.gov
Plate: LINC2091 row: P column: 02
High quality sequence scop: 634.
Location/Qualifiers
1..1011
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5923441"
/tissue_type="neuroblastoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 47"
/note="Organ: brain; Vector: pOTB7, site 1: XhoI, site 2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adapter: GGCACGAG(G). Size-selected >500bp for average
insert size 1.8kb. Library constructed by Ling Hong in
the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."

ORIGIN

Query Match 14.8%; Score 913.6; DB 5; Length 1011;
 Best Local Similarity 97.8%; Pred. No. 1.7e-228;
 Matches 945; Conservative 0; Mismatches 18; Indels 3; Gaps 2;

2701 CGGGAGGTCAGCTCTAGAGACGAGAGCGAAACTGAGACTCAAGGCGCAGCTCACAGAG 2760
 1 CGGGAGGTCAGCTCTAGAGACGAGAGCGAAACTGAGACTCAAGGCGCAGCTCACAGAG 60
 2761 CTACAGCTCTCCCTGAGAGGCGCAGTCAAGTTGACAGCCCTGAGGCTGACAGGCG 2820
 61 CTACAGCTCTCCCTGAGAGGCGCAGTCAAGTTGACAGCCCTGAGGCTGACAGGCG 120
 2821 GCCCTGAGAGGCGCAGCTTGGCCAGGAGACAGAGCTGAAAGACCAAGAGAGCT 2880
 121 GCCCTGAGAGGCGCAGCTTGGCCAGGAGACAGAGCTGAAAGACCAAGAGAGCT 180
 2881 GAAAGAGAGATCCAGGCACTCAGGCACTAGAGATGAAATCCAGGCGAAATTTGATGCT 2940
 181 GAAAGAGAGATCCAGGCACTCAGGCACTAGAGATGAAATCCAGGCGAAATTTGATGCT 240
 2941 CTTGCTAAGCTGTAAGTCTGTAATCAAGACTGAGAGAGAGCTTAAACAGCTGACCGAG 3000
 241 CTTGCTAAGCTGTAAGTCTGTAATCAAGACTGAGAGAGAGCTTAAACAGCTGACCGAG 300
 3001 GACAAAGCTGAAGTCAAGCAACCAAACTTCTAAGTCTGCAAACTGATGAGGCTTCT 3060
 301 GACAAAGCTGAAGTCAAGCAACCAAACTTCTAAGTCTGCAAACTGATGAGGCTTCT 360
 3061 GGGCCCAAGAGAGATTTGTATCAACTGCGAAGTGAAGTGAACATTTCCGCGGAGATC 3120
 361 GGGCCCAAGAGAGATTTGTATCAACTGCGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 420
 3121 ACGGAAAGAGAGATGAGCTTACAGCAGAGAGCAAGAGAGAGAGCTCTGAAGAGCAG 3180
 421 ACGGAAAGAGAGATGAGCTTACAGCAGAGAGCAAGAGAGAGAGCTCTGAAGAGCAG 480
 3181 TGCAACATCTGAGAGAGAGAGTCAATGATTTGAGAGGCTTAAACATGAGCTGCTAGAA 3240
 481 TGCAACATCTGAGAGAGAGAGTCAATGATTTGAGAGGCTTAAACATGAGCTGCTAGAA 540
 3241 AAAAGAGAGAGAGAGAGAGCTTGAAGAGAGAGCTTGAAGAGAGAGAGAGAGAGAGAG 3300
 541 AAAAGAGAGAGAGAGAGAGCTTGAAGAGAGAGAGCTTGAAGAGAGAGAGAGAGAGAG 600
 3301 TGTCGGGTTGAGAGCTGAGAGAGAGAGTGAACAGGAGAGAGAGAGAGAGAGAGAGAG 3360
 601 TGTCGGGTTGAGAGCTGAGAGAGAGAGTGAACAGGAGAGAGAGAGAGAGAGAGAGAG 660
 3361 GATCAGCGAGATCAACGAGTCTGCGCAGTGTGTGAGCTGGCAGTGAAGAGCAAGGCT 3420
 661 GATCAGCGAGATCAACGAGTCTGCGCAGTGTGTGAGCTGGCAGTGAAGAGCAAGGCT 720
 3421 GAGATTTCTGCTTGAAGAGAGAGCTTCAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3480
 721 GAGATTTCTGCTTGAAGAGAGAGCTTCAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 780
 3481 GACAGAGCTCAAG 3540
 781 GACAGAGCTCAAG 840
 3540 ACGAG 3599
 841 ACGAG 900
 3600 ATTACAG 3657
 901 ATTACAG 960
 3658 CAAGAA 3663
 961 CCAGAA 966

RESULT 2
 BQ228524 879 bp mRNA linear EST 02-MAY-2002
 LOCUS BQ228524
 DEFINITION AGENCOURT 7600932 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:6061485
 5', mRNA sequence.

ACCESSION BQ228524
 VERSION BQ228524.1 GI:20409924
 KEYWORDS EST.

SOURCE
 ORGANISM Homo sapiens (human)
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE
 AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: ATCC/DCTD/DTF

CDNA Library Preparation: Life Technologies, Inc.
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LLM1331 row: 9 column: 22
 High quality sequence stop: 636.
 Location/Qualifiers

FEATURES

1..879
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:6061485"
 /cuisine_type="melanotic melanoma"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH MGC 72"
 /note="Organ: skin; Vector: pCW-SPORT6; Site 1: NotI;
 Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.
 Average insert size 2 kb. Library constructed by Life
 Technologies."

ORIGIN

Query Match 13.3%; Score 822.6; DB 5; Length 879;
 Best Local Similarity 98.0%; Pred. No. 1.5e-204;
 Matches 853; Conservative 0; Mismatches 15; Indels 2; Gaps 2;

3949 CGGGAGGAGAGCTGCCACCGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 4008
 1 CGGGAGGAGAGCTGCCACCGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 60
 4009 GCGAGGCAAGAGATGCGCAATGTCGCGCATGTCGCGTCCAGAGAGCAAGCCAGTGCC 4068
 61 GCGAGGCAAGAGATGCGCAATGTCGCGCATGTCGCGTCCAGAGAGCAAGCCAGTGCC 120
 4069 ATGAGCTTCTGAGCGCCCGCATCCAGCCGCAAGAGAGAGAGAGAGAGAGAGAGAGAG 4128
 121 ATGAGCTTCTGAGCGCCCGCATCCAGCCGCAAGAGAGAGAGAGAGAGAGAGAGAGAG 180
 4129 AGTCGGCGCTTAAG 4188
 181 AGTCGGCGCTTAAG 240
 4189 AACATGCAAGCAAG 4248
 241 AACATGCAAGCAAG 300
 4249 TCCAATGCTCTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4308
 301 TCCAATGCTCTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 360
 4309 ACTGCGGCTTGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4368
 361 ACTGCGGCTTGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 420

QY 4369 ATGAACTCCCGAGTCTCCAGACCAAGAGCCAGCAGCTTGGACCTGGAAGGATG 4428
 |||||
 DB 421 ATGAACTCCCGAGTCTCCAGACCAAGAGCCAGCAGCTTGGACCTGGAAGGATG 480
 |||||
 QY 4429 ATGAAAGTGGCCAGAAATACAAAGAGGACAGAGGCTGGAGCAGAAATCATTTGC 4488
 |||||
 DB 481 ATGAAAGTGGCCAGAAATACAAAGAGGACAGAGGCTGGAGCAGAAATCATTTGC 540
 |||||
 QY 4489 CTGAGAGGATCAAAAGTCTCATTTATGACAAATGAAGCCAGAGAGCTGGACAGAGCCG 4548
 |||||
 DB 541 CTGAGAGGATCAAAAGTCTCATTTATGACAAATGAAGCCAGAGAGCTGGACAGAGCCG 600
 |||||
 QY 4549 GTGGAAGATTGAGCTGTGCTCTCCGAGCGGAGATGATCTATTCATGCTGCTTGGT 4608
 |||||
 DB 601 GTGGAAGATTGAGCTGTGCTCTCCGAGCGGAGATGATCTATTCATGCTGCTTGGT 660
 |||||
 QY 4609 GCTTCGGAATCGCAATACAGCCAAAGAGAGATGCTCCATACATCTGAAAGATGAATCT 4668
 |||||
 DB 661 GCTTCGGAATCGCAATACAGCCAAAGAGAGATGCTCCATACATCTGAAAGATGAATCT 720
 |||||
 QY 4669 CACCCGACACCACTGCT-GGCCCGGAGAACCTCTACTTCTAGCTCCAGCTTCCG 4727
 |||||
 DB 721 CAAACGCAACCACTGCTGCTGGGCCCGGAGAACCTCTACTTCTAGCTCCAGCTTCCG 780
 |||||
 QY 4728 TGAACAACAGCGCTGGGTCAACGCGCTTAAGATGATGTCAGAGT-GGAGAGATTCTA 4786
 |||||
 DB 781 TGAACAACAGCGCTGGGTCAACGCGCTTAAGATGATGTCAGAGTGGGAGAGATTCTA 840
 |||||
 QY 4787 GGGAAAAGCAAGAGCTGATGCTAACTGC 4816
 |||||
 DB 841 GGGAAAAGCAAGAGCTGATGCTTGGCCGC 870
 |||||

RESULT 3
 B0071141
 LOCUS B0071141 956 bp mRNA 1linear EST 02-APR-2002
 DEFINITION AGENCOURT 6853098 NIH_MGC_47 Homo sapiens cDNA clone IMAGE:5927502
 5', mRNA sequence..
 B0071141
 ACCESSION B0071141 GI:19900187
 VERSION B0071141
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 1 (bases 1 to 956)
 NIH-MGC http://mgi.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgaabs-remail.nih.gov
 Tissue Procurement: ATCC
 cDNA Library Preparation: Rubin Laboratory
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/ILNL at:
 http://image.jnl.gov
 Place: LNC2102 row: 1 column: 07
 High quality sequence scop: 650.
 Location/Qualifiers
 1..956
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5927502"
 /issue_type="neuroblastoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /clone_id="NIH_MGC_47"
 /note="Organ: brain; Vector: pOTB7, Site_1: XhoI, Site_2:
 EcoRI; cDNA made by oligo-dT priming. Directionally
 cloned into EcoRI/XhoI sites using the following 5'

adaptor: GGCAACGAG(G). Size-selected >500bp for average
 insert size 1.8kb. Library constructed by Ling Hong in
 the laboratory of Gerald M. Rubin (University of
 California, Berkeley) using ZAP-cDNA synthesis kit
 (Stratagene) and Superscript II RT (Life Technologies).
 Note: this is a NIH_MGC Library."

Query Match 13.2%; Score 814; DB 5; Length 956;
 Best Local Similarity 97.0%; Pred. No. 2.9e-202;
 Matches 873; Conservative 0; Mismatches 20; Indels 7; Gaps 4;

ORIGIN
 QY 2701 CCGAGAGTCACTCTAGAGCAAGAGAGCAAGAACTGAGCTCAAGCGCCAGCTCAAGAG 2760
 |||||
 DB 1 CCGAGAGTCACTCTAGAGCAAGAGAGCAAGAACTGAGCTCAAGCGCCAGCTCAAGAG 60
 |||||
 QY 2761 CTACAGCTCTCCTGAGAGAGCGGAGTCAAGCTTGAACGCTTGAAGCTGACCGGCG 2820
 |||||
 DB 61 CTACAGCTCTCCTGAGAGAGCGGAGTCAAGCTTGAACGCTTGAAGCTGACCGGCG 120
 |||||
 QY 2821 GCCCTGAGAGCAGCTTGCAGAGCGGAGAGCAAGAGCTGAAAGAGACACAGAGAGCT 2880
 |||||
 DB 121 GCCCTGAGAGCAGCTTGCAGAGCGGAGAGCAAGAGCTGAAAGAGACACAGAGAGCT 180
 |||||
 QY 2881 GAAGAGAGATCAAGCACTCAAGCAATAGAGATGAATCAAGCGCAATTTGATGCT 2940
 |||||
 DB 181 GAAGAGAGATCAAGCACTCAAGCAATAGAGATGAATCAAGCGCAATTTGATGCT 240
 |||||
 QY 2941 CTTCTGAACGCTGATCTGATTAATCAAGAGCTGAGAGAGACAGTAAACAGCTGACCGAG 3000
 |||||
 DB 241 CTTCTGAACGCTGATCTGATTAATCAAGAGCTGAGAGAGACAGTAAACAGCTGACCGAG 300
 |||||
 QY 3001 GACAAGCTGAAGTCAACCAACCAAACTCTACTTGTCCAAACCAATCTGATGAGCTTCT 3060
 |||||
 DB 301 GACAAGCTGAAGTCAACCAACCAAACTCTACTTGTCCAAACCAATCTGATGAGCTTCT 360
 |||||
 QY 3061 GGCGCCCAACGAGAGATGTAACACTGCAAGTGAAGTGAACATCTCCGCGGAGATC 3120
 |||||
 DB 361 GGCGCCCAACGAGAGATGTAACACTGCAAGTGAAGTGAACATCTCCGCGGAGATC 420
 |||||
 QY 3121 ACCGAAACGAGATGAGCTTACCAAGCCAGAAAGCAAAACGATGAGAGCTTGAAGACAG 3180
 |||||
 DB 421 ACCGAAACGAGATGAGCTTACCAAGCCAGAAAGCAAAACGATGAGAGCTTGAAGACAG 480
 |||||
 QY 3181 TGCACATCTGAGAGAAAGAGTCAATGATTTGAGGCGCTTAACGATGAGCTGAGAA 3240
 |||||
 DB 481 TGCACATCTGAGAGAAAGAGTCAATGATTTGAGGCGCTTAACGATGAGCTGAGAA 540
 |||||
 QY 3241 AAAAGACGAGAGTGGAGAGCTGAGAGAGCGTCTGAGTGAAGTGAATCCAGATTGAG 3300
 |||||
 DB 541 AAAAGACGAGAGTGGAGAGCGTCTGAGTGAAGTGAATCCAGATTGAG 600
 |||||
 QY 3301 TGTGCGGTTGAGAGCTGAGAGATGCTGACACCGAGAAACAGAGCAGGCGAGAGCC 3360
 |||||
 DB 601 TGTGCGGTTGAGAGCTGAGAGATGCTGACACCGAGAAACAGAGCAGGCGAGAGCC 660
 |||||
 QY 3361 GATCAGCGGATACCGAGTCTGCCAGAGTGTGAGAGCTGSCAGTGAAGAGACAGAGCT 3420
 |||||
 DB 661 GATCAGCGGATACCGAGTCTGCCAGAGTGTGAGAGCTGSCAGTGAAGAGACAGAGCT 720
 |||||
 QY 3421 GAGATTCTGCTCTGAGAGAGCTCTCAAGAGACAGAACTGAAGCCGAGAGCTCTCT 3480
 |||||
 DB 721 GAGATTCTGCTCTGAGAGAGCTCTCAAGAGACAGAACTGAAGCCGAGAGCTCTCT 780
 |||||
 QY 3481 GACAGCTCAATGACCT-GGAGAGAGAGATGCTTGA--TGAATGCCGAGAC 3537
 |||||
 DB 781 GACAGCTCAATGACCTGGAGAGAGAGATGCTTGA--TGAATGCCGAGAC 840
 |||||
 QY 3538 TTAACAGAGAA--GGTGAAGACTGAAC--GAGAGCTCAACAGAGGCTTCTGAAGAGCA 3593
 |||||
 DB 841 TTAACAGAGAAAGCTGGAGAGTGAACGAGAAAGCTCAACAGAGGCTTCTGGAGAGCA 900
 |||||

```

RESULT 4
LOCUS      BU181633                830 bp      mRNA      linear      EST 04-SEP-2002
DEFINITION AGNCOURT_7906225 NIH_MGC_67 Homo sapiens cDNA clone IMAGE:6140538
            5', mRNA sequence.
ACCESSION  BU181633
VERSION    BU181633.1  GI:22695617
KEYWORDS   EST.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE  1 (bases 1 to 830)
            NIH-MGC http://mgs.nci.nih.gov/.
            National Institutes of Health, Mammalian Gene Collection (MGC)
            Unpublished (1999)
COMMENT    Contact: Robert Strausberg, Ph.D.
            Email: cgabbs-r@mail.nih.gov
            Tissue Procurement: ATCC
            cDNA library Preparation: Life Technologies, Inc.
            DNA Sequencing by: Agencourt Bioscience Corporation
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: L1AM13459 row: e column: 19
            High quality sequence stop: 652.
FEATURES
    source
        1..830
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="IMAGE:6140538"
            /issue_type="retinoblastoma"
            /lab_host="DH10B (phage-resistant)"
            /note="Organ: eye; Vector: pCMV-SPORT6; Site 1: NotI;
            Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.
            Average insert size 1.75 kb. Library constructed by Life
            Technologies."
ORIGIN
Query Match      13.1%; Score 810.2; DB 5; Length 830;
Best Local Similarity 99.4%; Pred. No. 2.7e-201;
Matches 823; Conservative 0; Mismatches 4; Indels 1; Gaps 1;

QY      2435 TGGATTCCAGATCAATCCCTTGAAACAGAGATTGTGGAACCTGTGTAAGCCAAATTAAC 2494
DB      1 TGGATTCCAGATCAATCCCTTGAAACAGAGATTGTGGAACCTGTGTAAGCCAAATTAAC 60

QY      2495 TTGCAGCAATATAGACGCTTTTACCAAGAAAGAAAGAGAGATGATTT 2554
DB      61 TTGCAGCAATATAGACGCTTTTACCAAGAAAGAAAGAGAGATGATTT 120

QY      2555 CTGAACCTCAGGCAAGAAATTTTACCTGAGACACAGGCTGGAAAGTTGAGAGCCCAAG 2614
DB      121 CTGAACCTCAGGCAAGAAATTTTACCTGAGACACAGGCTGGAAAGTTGAGAGCCCAAG 180

QY      2615 ACCGAAATCTGAGAGACGAGCTGGAAGATCGCCACCAAGCCACAGTGAAGAATTC 2674
DB      181 ACCGAAATCTGAGAGACGAGCTGGAAGATCGCCACCAAGCCACAGTGAAGAATTC 240

QY      2675 GGCTGTGGAACCTGAGACAAAGATTGCGGAGGCTGAGTCTAGAGACAGAGAGCAGAAAC 2734
DB      241 GGCTGTGGAACCTGAGACAAAGATTGCGGAGGCTGAGTCTAGAGACAGAGAGCAGAAAC 300

QY      2735 TGGAGCTCAAGCGCCAGCTCAAGAGCTACAGCTCTCCCTGACAGAGCGGAGTCAAGT 2794
DB      301 TGGAGCTCAAGCGCCAGCTCAAGAGCTACAGCTCTCCCTGACAGAGCGGAGTCAAGT 360

QY      2795 TGAACAGCCCTGACAGGCTGACAGGCGGCGCTTGGAGAGACCACTTCCGACAGCGAAGACG 2854
DB      361 TGAACAGCCCTGACAGGCTGACAGGCGGCGCTTGGAGAGACCACTTCCGACAGCGAAGACG 420

```

```

QY      2855 AGCTGAAAGAGACCAACAGCAAGAGCTGAAGAGAGATCCAGGACTCATGCGCATATGAG 2914
DB      421 AGCTGAAAGAGACCAACAGCAAGAGCTGAAGAGAGATCCAGGACTCATGCGCATATGAG 480

QY      2915 ATGAAATTCAGGCGCAATTTGATGCTTCGTAACAGAGTGTCTGTAATCAAGAGACTG 2974
DB      481 ATGAAATTCAGGCGCAATTTGATGCTTCGTAACAGAGTGTCTGTAATCAAGAGACTG 540

QY      2975 AGAGAGCTTAACACAGCTGACCGAGAGCAACGCTGAATCAACACCAAACTTCTACT 3034
DB      541 AGAGAGCTTAACACAGCTGACCGAGAGCAACGCTGAATCAACACCAAACTTCTACT 600

QY      3035 TGTCCAAACACTCGATGAGGCTTTGCGGCCAAGACAGAGATTGTATCAACTGCGAAGT 3094
DB      601 TGTCCAAACACTCGATGAGGCTTTGCGGCCAAGACAGAGATTGTATCAACTGCGAAGT 660

QY      3095 AAGTGACCATCTCCGCGCGGAGATCAAGGACAGAGATGAGCTTACAGCCAGAAC 3154
DB      661 AAGTGACCATCTCCGCGCGGAGATCAAGGACAGAGATGAGCTTACAGCCAGAAC 720

QY      3155 AAACAGTGAAGGCTCTGAAGACCAAGTGCACCATGCTGAGAGAACAGTCAATGATTT-G 3213
DB      721 AAACAGTGAAGGCTCTGAAGACCAAGTGCACCATGCTGAGAGAACAGTCAATGATTTGG 780

QY      3214 GAGGCCCTTAACGATGAGCTGCTAAGAAAAGAGCGGCACTGAGAGCC 3261
DB      781 GAGGCCCTTAACGATGAGCTGCTAAGAAAAGAGCGGCACTGAGAGCC 828

RESULT 5
LOCUS      BI253509                849 bp      mRNA      linear      EST 17-JUL-2001
DEFINITION 602973370F1 NIH_MGC_12 Homo sapiens cDNA clone IMAGE:5112737 5',
            mRNA sequence.
ACCESSION  BI253509
VERSION    BI253509.1  GI:14805003
KEYWORDS   EST.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE  1 (bases 1 to 849)
            NIH-MGC http://mgs.nci.nih.gov/.
            National Institutes of Health, Mammalian Gene Collection (MGC)
            Unpublished (1999)
COMMENT    Contact: Robert Strausberg, Ph.D.
            Email: cgabbs-r@mail.nih.gov
            Tissue Procurement: ATCC
            cDNA library Preparation: Life Technologies, Inc.
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: L1AM11273 row: 1 column: 18
            High quality sequence start: 2
            High quality sequence stop: 772.
FEATURES
    source
        1..849
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="IMAGE:5112737"
            /issue_type="cervical carcinoma cell line"
            /lab_host="DH10B"
            /clone_11b="NIH_MGC_12"
            /note="Organ: cervix; Vector: pCMV-SPORT6; Site 1: NotI;
            Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.
            Average insert size 1.4 kb. Library prepared by Life
            Technologies."
ORIGIN

```


Db 540 CCGCGAGGGGCGGACCGAGCTGGCGAGGACAAAGTCTCTGGCGCGCCCTCGAGAGGAA 599
Qy 6018 GAAATCCCCCGGAGTACTAGCAGCGGAGAGAGCGAGTCCCGCCGAGGCTGTTGA 6077
Db 600 GAAATCCCCCGGAGTACTAGCAGCGGAGAGAGCGAGTCCCGCCGAGGCTGTTGA 659
Qy 6078 AGACAGCAGCAGAGGGGCGGCTGCTCGGAGAGCGCTGAGAGCCCGCTGTCCAGGTGAA 6137
Db 660 AACAAGCAGCAGAGGGGCGGCTGCTCGGAGAGCGCTGAGAGCCCGCTGTCCAGGTGAA 719
Qy 6138 CAAAGTCTGGGACCAAGTCTTCAGATATA 6165
Db 720 CAAAGTCTGGGACCAAGTCTTCAGATATA 747
RESULT 7
CN281134
LOCUS 1700423750642 GRN_EB Homo sapiens cDNA 5', mRNA EST 16-MAY-2004
DEFINITION CN281134
ACCESSION CN281134
VERSION CN281134.1 GI:47297548
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 757)
Brandenberger, R., Wei, H., Zhang, S., Lei, S., Murgue, J., Fisk, G. J.,
Li, Y., Xu, C., Fang, R., Guegler, K., Rao, W. S., Mandalam, R.,
Lebkowaki, J and Stanton, L. W.
Transcriptome characterization elucidates signaling networks that
control human ES cell growth and differentiation
Nat. Biotechnol. 22 (6), 707-716 (2004)
CONTACT: Brandenberger R
Regenerative Medicine
Genon Corporation
230 Constitution Drive, Menlo Park, CA 94025, USA
Tel: 650 473 8658
Fax: 650 473 7760
Email: rbrandenberger@genon.com
Insert Length: 757 Std Error: 0.00.
Location/Qualifiers
1. 757
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/feature_type="embryonic stem cells, embryoid bodies"
derived from H1, H7 and H9 cells"
/clone_1lb="GRN EB"
/note="oligo dt primed, full-length enriched cDNA library
from embryoid body outgrowth derived from h9 cell lines
H1 (p32), H7 (p29), and H9 (p26) maintained in feeder-free
conditions."

ORIGIN
Query Match 11.9%; Score 732; DB 7; Length 757;
Best Local Similarity 99.7%; Pred. No. 1e-180; 0; Indels 2; Gaps 2;
Matches 754; Conservative 0; Mismatches 0;
Qy 1551 AAAGCGAATTGGAGCAGCAGATGAGGTGTCCAGAGAGATGACAAAGCACTGCA 1610
Db 3 AAAGCGAATTGGAGCAGCAGATGAGGTGTCCAGAGAGATGACAAAGCACTGCA 62
Qy 1611 GCTTCTCATGATATCAGAGAGCAGCGGAAAGCTTCAAGAAATCAAAGAGCAGAGTA 1670
Db 63 GCTTCTCATGATATCAGAGAGCAGCGGAAAGCTTCAAGAAATCAAAGAGCAGAGTA 122
Qy 1671 CCAAGGCTCAAGTGAAGAAATGAGTTGATGATGAAATCAAGTTGGAAGAGATCTTGTCTC 1730
Db 123 CCAAGGCTCAAGTGAAGAAATGAGTTGATGATGAAATCAAGTTGGAAGAGATCTTGTCTC 182
Qy 1731 AGCAAGAGACGAGATGATCTTACGAATCTGAGCTGAGAGACTCTCGGCTTGCTGCTGA 1790

Db 183 AGCAAGAGACGAGATGATCTTACGAATCTGAGCTGAGAGACTCTCGGCTTGCTGCTGA 242
Qy 1791 AGAATTTCAAGCGGAAAGGACAGAAATGTCAAGTAAATCTGTGAAAGCTTAAGATCAAG 1850
Db 243 AGAATTTCAAGCGGAAAGGACAGAAATGTCAAGTAAATCTGTGAAAGCTTAAGATCAAG 302
Qy 1851 GAAGCTTGAAGTGGAGAAATATGCGAAACTGGAGAAATCAATGCTGAGACAGACTCA 1910
Db 303 GAAGCTTGAAGTGGAGAAATATGCGAAACTGGAGAAATCAATGCTGAGACAGACTCA 362
Qy 1911 AATTAGAGAGCTCCAA-GAGAACTGGAAGGCTGTAAAGCCAGCAAGGAGCCACCG 1969
Db 363 AATTAGAGAGCTCCAAAGAGAACTGAGAAAGGCTGTAAAGCCAGCAAGGAGCCACCG 422
Qy 1970 AGCTGCTGAGAAATTCGCGCAGGAGAGCGGAGCGGAGAGGAGCTGAGAAAGCTGC 2029
Db 423 AGCTGCTGAGAAATTCGCGCAGGAGAGCGGAGCGGAGAGGAGCTGAGAAAGCTGC 482
Qy 2030 AGAACCGAGAGGATTTCTTGAAGGCATCAGAAAGAGTGGTGAAGCTGAGAAAGCTGC 2089
Db 483 AGAACCGAGAGGATTTCTTGAAGGCATCAGAAAGAGTGGTGAAGCTGAGAAAGCTGC 542
Qy 2090 GCCATTCTTGAGAGAAAGGTAAGAGACTGAGACATGAGGGTGAAGAAACAGAC 2149
Db 543 GCCATTCTTGAGAGAAAGGTAAGAGACTGAGACATGAGGGTGAAGAAACAGAC 602
Qy 2150 TGAAGATGACATCCAGACAAATATCCCAAGATCCAGAGATGAGCTGATAAATTTCTGG 2209
Db 603 TGAAGATGACATCCAGACAAATATCCCAAGATCCAGAGATGAGCTGATAAATTTCTGG 662
Qy 2210 AGCTGGAAGAGAAACATGGGAGGCGCCAGAGTTCAGGCCAGCACTTAAGATGACACTGA 2269
Db 663 AGCTGGAAGAGAAACATGGGAGGCGCCAGAGTTCAGGCCAGCACTTAAGATGACACTGA 721
Qy 2270 AACAGAAAGAGAGACATATGAGAAAGATTAAAG 2305
Db 722 AACAGAAAGAGAGACATATGAGAAAGATTAAAG 757
RESULT 8
LOCUS BC031156 958 bp mRNA linear HTC 04-MAR-2003
DEFINITION Mus musculus, Similar to citron, clone IMAGE:4976752, mRNA.
ACCESSION BC031156
VERSION BC031156.1 GI:21411076
KEYWORDS HTC.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 958)
Strausberg, R.
Direct Submission
Submitted (03-JUN-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: <http://mgc.nci.nih.gov>
CONTACT: MGC help desk
Email: cgapbs-remail.nih.gov
Tissue Procurement: Jeffrey Green M.D.
cDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Sequencing Group at the Stanford Human Genome
Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: <http://www-shgc.stanford.edu>
Contact: (Dickson, Mark) mcdepaxil.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
R. M.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/BLM at: <http://image.llnl.gov>

Series: IRAK Plate: 59 Row: J Column: 6
This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis

FEATURES
This clone has the following problem: retained intron.

source

1. 958
/organism="Mus musculus"
/mol_type="mRNA"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:4976752"
/issue_type="Mammary tumor, C3(1)-Tag model. Infiltrating ductal carcinoma. 5 month old virgin mouse."
/clone_lib="NCI_CGAP_Mam6"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"

ORIGIN

Query Match 11.4%; Score 705.6; DB 3; Length 958;
Best Local Similarity 89.1%; Pred. No. 1e-173;
Matches 775; Conservative 0; Mismatches 89; Indels 6; Gaps 1;

QY 5302 AATTACGATCTCTCAATTGGAACCAATTAATTTACGAATTCACATGAAAGCAGTACAG 5361
DB 40 AAGAAAGCAATCTCTCATGTGGACCAACAAATTTATGATGACATGAAAGCAGTACAG 99
QY 5362 CTCGAGAAATCTCTGATTAAGATGACATCTCTTGGACCTCTGTGTTTGGCCCTCT 5421
DB 100 CTGTGATGAGTCTCTGACAAAGAACGACATCTTGGACCTCTGTGTTTGGCCCTCTG 159
QY 5422 TCCACAGCTTCCCTCTCTCAATCTGTGACAGTGAACGCGCAGGAGGAGAGAGATAC 5481
DB 160 TCCACAGCTTCCCTCTCTCTCAATCTGTGACAGGCGACAGCGCGGAGAGAGAGATAC 219
QY 5482 TTGCTGTGTTTCCAGAAATTTGAGTGTCTGTGATTTCTTACGGAAGACTACCCGACA 5541
DB 220 CTGCTGTGTTTCCAGAAATTTGAGTGTCTGTGATTTCTTACGGAAGACTACCCGACA 279
QY 5542 GACGATCTCAAGTGAAGTGTGCTTACCTTTGGCTTACAGAAACCTATCTGTTT 5601
DB 280 GATGATCTTAAGTGAAGTGTGCTTACCTTGGCTTACAGAAACCTATCTGTTT 339
QY 5602 GTGACCCACTTCACTCACTCGAAGTAAATTGATGATCAGGACGCTCTTACAGAGGAGC 5661
DB 340 GTGACCTCACTTCACTCTCTGGAAGTCAATTGATGATCAGGACGCTCTTACAGAGGAGC 399
QY 5662 CCTGCCGAGCGGATCTGTGACATCCGAAACCGCGCTTACCTGGGCTTGCATTTCTTA 5721
DB 400 CCTGCCGAGCGGATCTGTGAAATTTCCAAACCTCTGCTACTGGGCTTGCATTTCTTC 459
QY 5722 GGAGGATTTTACTGTGCGTCTCTCATACAGAAATTAAGGTCATTTGCTGCAGAGGA 5781
DB 460 GGAGGATTTTACTGTGCGTCTCTCATACAGAAATTAAGGTCATTTGCTGCAGAGGA 519
QY 5782 AACCTGTGAAGAGTCCGAGCATGAACCAACCGGAGCGCTGTCACCTCCGACAGCAG 5841
DB 520 AACCTGTGAAGAGTCCGAGCATGAACCAACCGGAGCGCTGTCACCTCCGACAGCAG 579
QY 5842 CCCAACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 5901
DB 580 CCCAACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 639
QY 5902 CCAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 5959
DB 640 CCAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 699
QY 5960 ----GCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 6015
DB 700 GACAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 759
QY 6016 GAGGAATCCCCGCGGAGTACTCAGACAGCGGAGAGAGGAGGAGGAGGAGGAGGAGGAGG 6075

DB 760 GAGAAATCCCCAGGCGGATGCTCAGACATGAGAGAGGCGTCCCGAGGAGACTGTTT 819
QY 6076 GAAGACAGCAGCAGGAGGCGGCTGTCTGCGGAGCCGTGAGAGCCCGCTGTCCAGATG 8135
DB 820 GAAGACAGCAGCAGGAGGCGGCTGTCTGCGGAGCCGTGAGAGCCCGCTGTCCAGATG 879
QY 6136 AACAAAGTCTGGGAGCAACAGTCTTCAATATA 6165
DB 880 AACAAAGTCTGGGAGCAACAGTCTTCAATATA 909

RESULT 9

BG912161 843 bp mRNA linear EST 05-JUN-2001
LOCUS 602812833F1 NCI CGAP_Brn67 Homo sapiens cDNA clone IMAGE:4944657
DEFINITION 5', mRNA sequence.

ACCESSION BG912161 GI:114292637
VERSION BG912161.1

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

1 (bases 1 to 843)
NIH-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: David N. Louis, M.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:
<http://image.llnl.gov>
Plate: L1AM10890 row: a column: 10
High quality sequence stop: 778.

FEATURES

source

1. 843
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4944657"
/issue_type="anaplastic oligodendroglioma with 1p/19q loss"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NCI_CGAP_Brn67"
/note="Organ: brain; Vector: pCMV-SPORT6; Site 1: NCI; Site 2: Salt; Cloned unidirectionally. Primer: oligo dT. Average insert size 2.3 kb. Constructed by Life Technologies. Note: this is a NCI_CGAP Library."

ORIGIN

Query Match 11.3%; Score 698.2; DB 4; Length 843;
Best Local Similarity 93.3%; Pred. No. 8.5e-172;
Matches 785; Conservative 0; Mismatches 48; Indels 8; Gaps 5;

QY 4669 CACCGGACACCACTGCTGCGCCCGGAGAACCTCTACTTGTAGCTCCAGCTTCCT 4728
DB 2 CACGGGTGACCACTGCTGCGCCCGGAGAACCTCTACTTGTAGCTCCAGCTTCCT 61
QY 4729 GACAAACAGCGTGGGTGACCGGCTTAGAATAGTGTGCGAGGTGGAGAGTTCTAG 4788
DB 62 GACAAACAGCGTGGGTGACCGGCTTAGAATAGTGTGCGAGGTGGAGAGTTCTAG 121
QY 4789 GAAAAAGCAGAGTGAATCTGCTTGAATCTCCCTGCTGAATCTGGAAGGTGAT 4848
DB 122 GAAAAAGCAGAGTGAATCTGCTTGAATCTCCCTGCTGAATCTGGAAGGTGAT 181
QY 4849 GACCGTTAGACATGAATGACGCGCTTCACTGACGAGGTGTGTGGGACCC 4908
DB 182 GACCGTTAGACATGAATGACGCGCTTCACTGACGAGGTGTGTGGGACCC 241

QY 4909 GAGGAGGGCTCTAGCCCTGGAATGCTTGAAGAACTCCCTAACCCATGTCCAGGAATT 4968
| | | | |
DB 242 GAGGAGGGCTCTAGCCCTGGAATGCTTGAAGAACTCCCTAACCCATGTCCAGGAATT 301
| | | | |
QY 4969 GAGGAGCTTTTCCAAATTTATATATATCAAGAGCTTGAAGAGAGTACTATGATAGAGAA 5028
| | | | |
DB 302 GAGGAGCTTTTCCAAATTTATATATATCAAGAGAGCTTGAAGAGAGTACTATGATAGAGAA 361
| | | | |
QY 5029 GAAGAGCGGGCACTGCTGTCTTGTGAAGTGAAGAAAGTGAACAGTCCCTGGCCAGTCC 5088
| | | | |
DB 362 GAAGAGCGGGCACTGCTGTCTTGTGAAGTGAAGAAAGTGAACAGTCCCTGGCCAGTCC 421
| | | | |
QY 5089 CACCTGCTGCGCCAGCCGCAATTCACCCCAATTTTGAAGCTGTCAAAGGCTGCGCA 5148
| | | | |
DB 422 CACCTGCTGCGCCAGCCGCAATTCACCCCAATTTTGAAGCTGTCAAAGGCTGCGCA 481
| | | | |
QY 5149 TTGTTTGGGGGAGGCAAGTTGAGAAAGGAGCTTGTGATCTGTGCAAGCCATGCCCCAGCAA 5208
| | | | |
DB 482 TTGTTTGGGGGAGGCAAGTTGAGAAAGGAGCTTGTGATCTGTGCAAGCCATGCCCCAGCAA 541
| | | | |
QY 5209 GTGCTCATCTCCGCTACAGCAAGAACTCAGCAAAATATGCAATCCGAAAGAGATAGAG 5268
| | | | |
DB 542 GTGCTCATCTCCGCTACAGCAAGAACTCAGCAAAATATGCAATCCGAAAGAGATAGAG 601
| | | | |
QY 5269 ACCTGAGAGCCCTGAGCTGTATCCACTTACCAATTAAGATATCTCATT--GGAACCA 5327
| | | | |
DB 602 ACCTGAGAGCCCTGAGCTGTATCCACTTACCAATTAAGATATCTCATTGGGAAACCA 661
| | | | |
QY 5328 TAAATTTCTACG-AAATGACATGAA-GCAGTACACCTGAGAGAAAT--CTGTGATTAAGA 5383
| | | | |
DB 662 TAAATTTCTACGAAATGACATGAAAGGAGTACACCTGAGAGAAATTTCTGTGATTAAGA 721
| | | | |
QY 5384 ATGACCATCTCTTGACACCTGCTGTG---TTTGCGCGCTCTTCCAAAGCTTCCCTGTCT 5440
| | | | |
DB 722 ATGACCATCTCTTGACACCTGCTGTG---TTTGCGCGCTCTTCCAAAGCTTCCCTGTCT 781
| | | | |
QY 5441 CAATCTGAGAGTGAACAGCGAGGAGGAGAGAGTACTGTGTGTTTCCAGCAAT 5500
| | | | |
DB 782 AATCTGAGAGTGAACAGCGAGGAGGAGAGAGTACTGTGTGTTTCCAGCAAT 841
| | | | |
QY 5501 T 5501
| | | | |
DB 842 T 842
| | | | |
RESULT 10
CN295302 653 bp mRNA linear EST 16-MAY-2004
LOCUS 17000600183004 GRN_PRENEM Homo sapiens cDNA 5', mRNA sequence.
DEFINITION CN295302
ACCESSION CN295302
VERSION CN295302.1 GI:47311716
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 653)
Brandenberger, R., Wei, H., Zhang, S., Lei, S., Muzge, J., Fisk, G., J.,
Li, Y., Xu, C., Pang, R., Guegler, K., Rao, M. S., Mandalam, R.,
Lebkowski, J. and Stanton, L. W.
Transcriptome characterization elucidates signaling networks that
control human ES cell growth and differentiation
Nat. Biotechnol. 22 (6), 707-716 (2004)
CONTACT: Brandenberger R
Regenerative Medicine
Geon Corporation
230 Constitution Drive, Menlo Park, CA 94025, USA
Tel: 650 473 8658
Fax: 650 473 7760
Email: rbrandenberger@geon.com
Insert Length: 653 Std Error: 0.00.
Location/Qualifiers

source 1. .653
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/tisue_type="embryonic stem cell", retinoic acid and
mitogen-treated hES cell line H7"
/clone_lib="GRN_PRENEM"
/note="Oligo dt primed, full-length enriched cDNA library
from hES cell line H7 (p29) maintained in feeder-free
conditions. Embryoid bodies were generated in the presence
of all-trans retinoic and mitogens."

Query Match 10.5%; Score 649.4; DB 7; Length 653;
Best Local Similarity 99.8%; Pred. No. 5.5e-159;
Matches 650; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2017 CTGAGGAAGCTGCAACCCGAGAGATTTCTTGAAGGCTATGAGAAAGCTGTGTGAA 2076
| | | | |
DB 3 CGGAGGAAGCTGCAACCCGAGAGATTTCTTGAAGGCTATGAGAAAGCTGTGTGAA 62
| | | | |
QY 2077 GCTGAGGAACCGCGCATTTCTGTGAGAACAGATTAAGAGACTAGAGCCATGAGCGT 2136
| | | | |
DB 63 GCTGAGGAACCGCGCATTTCTGTGAGAACAGATTAAGAGACTAGAGCCATGAGCGT 122
| | | | |
QY 2137 AGAGAAACAGACTGAAAGATGATCTCCAGACAAATTTCCACAGATCCAGAGATGGCT 2196
| | | | |
DB 123 AGAGAAACAGACTGAAAGATGATCTCCAGACAAATTTCCACAGATCCAGAGATGGCT 182
| | | | |
QY 2197 GATTAATTTCTGGAAGCTGGAAGAGAAACATGCGGAGGCCCAAGCTCAGCCAGACCTA 2256
| | | | |
DB 183 GATTAATTTCTGGAAGCTGGAAGAGAAACATGCGGAGGCCCAAGCTCAGCCAGACCTA 242
| | | | |
QY 2257 GAAGTGCACTTGAAACAGAAAGAGAGAGACTATGAGAAAGATTAAGTGTGAGCAAT 2316
| | | | |
DB 243 GAAGTGCACTTGAAACAGAAAGAGAGAGACTATGAGAAAGATTAAGTGTGAGCAAT 302
| | | | |
QY 2317 CAGATTAAGAAAGACCTGCTGACAGAGAGACACTGTGAGAAACATATGATGAGAGACCGAG 2376
| | | | |
DB 303 CAGATTAAGAAAGACCTGCTGACAGAGAGACACTGTGAGAAACATATGATGAGAGACCGAG 362
| | | | |
QY 2377 GAGGAGGCCCATGAGAAAGGCAAAATTTCTCAGCGAACAGAGGCGATGATCATGTATG 2436
| | | | |
DB 363 GAGGAGGCCCATGAGAAAGGCAAAATTTCTCAGCGAACAGAGGCGATGATCATGTATG 422
| | | | |
QY 2437 GATTCGAAGATGATGATCCCTGTGAGAGAGATTTGAACTGTCTGAAGCCATTAATCTT 2496
| | | | |
DB 423 GATTCGAAGATGATGATCCCTGTGAGAGAGATTTGAACTGTCTGAAGCCATTAATCTT 482
| | | | |
QY 2497 GAGGCAATGAGAGTCTTTTAAACCAAGAGAAATGAGAGGCCCAAGAGAGATGATTTCT 2556
| | | | |
DB 483 GAGGCAATGAGAGTCTTTTAAACCAAGAGAAATGAGAGGCCCAAGAGAGATGATTTCT 542
| | | | |
QY 2557 GAATCTAGGCAACAGAAATTTTAACTGTGAGACAGAGGCTGGAGTGGAGGCCCAAGAC 2616
| | | | |
DB 543 GAATCTAGGCAACAGAAATTTTAACTGTGAGACAGAGGCTGGAGTGGAGGCCCAAGAC 602
| | | | |
QY 2617 CGAAAACTGAGAGAGAGCTGTGAGAGAGATCAAGCCACAGACAGTATC 2667
| | | | |
DB 603 CGAAAACTGAGAGAGAGCTGTGAGAGAGATCAAGCCACAGACAGTATC 653
| | | | |

RESULT 11
BF905370/c 652 bp mRNA linear EST 18-JAN-2001
LOCUS BF905370/c
DEFINITION IL3-MT0267-261200-410-H07 MT0267 Homo sapiens cDNA, mRNA sequence.
ACCESSION BF905370
VERSION BF905370.1 GI:12296829
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 652)
 AUTHORS Dae Neco,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Bala,G.S., Simpson,D.H., Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare,M.J., Soares,F., Brentant,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.
 TITLE Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
 MEDLINE 20202663
 PUBMED 10737800
 COMMENT Contact: Simpson A.J.G.
 Laboratory of Cancer Genetics
 Ludwig Institute for Cancer Research
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
 Tel: +55-11-2704922
 Fax: +55-11-2707001
 Email: asimpson@ludwig.org.br
 This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
 (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=IL3&f2=IL3-WT0267-261200-410-H07&f3=2000-12-26&f4=1)
 Seq primer: puc 18 forward
 High quality sequence stop: 624.
 Location/Qualifiers
 1..652
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /dev_stage="Adult"
 /clone_lib="WT0267"
 /note="Organ: marrow; Vector: puc18; Site 1: Smal; Site 2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the puc 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

ORIGIN
 Query Match 10.2%; Score 629.2; DB 2; Length 652;
 Best Local Similarity 98.0%; Pred. No. 1.1e-153;
 Matches 637; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 2525 GGAACTGAAGGCCCAAGAGATGATTTCTGAATCTGAGCAACGAATTTTACTCTG 2584
 DB 652 GGAACTGAAGGCCCAAGAGATGATTTCTGAATCTGAGCAACGAATTTTACTCTG 593
 QY 2585 AGACACAGGCTGGAAGTTGAGAGCCAGAACCGAAACTGAGAGAGAGCTGAGAGAGA 2644
 DB 592 AACACAGGCTGGAAGTTGAGAGCCAGAACCGAAACTGAGAGAGAGCTGAGAGAGA 533
 QY 2645 TAGACCAACCAAGCCACAGTGAACAGATTCGGCTGCTGGAATCTGAGACAAGATTGCGGG 2704
 DB 532 TAGACCAACCAAGCCACAGTGAACAGATTCGGCTGCTGGAATCTGAGACAAGATTGCGGG 473
 QY 2705 AGGTAGTGTAGAGACGAGAGCAAGAACTGAGAGTCAAGGCCAGCTTCAAGAGACTTAC 2764
 DB 472 AGGTAGTGTAGAGACGAGAGCAAGAACTGAGAGTCAAGGCCAGCTTCAAGAGACTTAC 413
 QY 2765 AGCTCTCCCTGAGAGCGCGAGTCAAGTGTACAGCCCTGAGAGCTGAGCGAGCGGCC 2824
 DB 412 AGCTCTCCCTGAGAGCGCGAGTCAAGTGTACAGCCCTGAGAGCTGAGCGAGCGGCC 353
 QY 2825 TGGAGAGCCAGCTTGGCCAGCGGAAGACAGAGCTGAGAGACCAAGAGAGAGTGAAG 2884
 DB 352 TGGAGAGCCAGCTTGGCCAGCGGAAGACAGAGCTGAGAGACCAAGAGAGAGTGAAG 293
 QY 2885 AGGAGATCCAGGACATCAGCGGACATAGAGATGAATCCAGGCGCAATTTGATGCTCTTC 2944
 DB 292 AGGAGATCCAGGACATCAGCGGACATAGAGATGAATCCAGGCGCAATTTGATGCTCTTC 233

QY 2945 GTAAACGCTGTACTGTAAATCAGACCTGTGAGAGAGCACTTAAACAGCTGACCGAGACA 3004
 DB 232 GTAAACGCTGTACTGTAAATCAGACCTGTGAGAGAGCACTTAAACAGCTGACCGAGACA 173
 QY 3005 AGCCTAATCTCAACCAACCAAACTTTTACTTTGTTCACCAACCACTCGATGAGCTTTCGGC 3064
 DB 172 AGCCTAATCTCAACCAACCAAACTTTTACTTTGTTCACCAACCACTCGATGAGCTTTCGGC 113
 QY 3065 CCACGACGAGATTTGTACAACTGCAAGTGAAGTGAACCATCTCCGCGGAGATCAACG 3124
 DB 112 CCACGACGAGATTTGTACAACTGCAAGTGAAGTGAACCATCTCCGCGGAGATCAACG 53
 QY 3125 AACGAGATGACGCTTACCAAGCCAGAACCAACGATGAGAGCTTGAAG 3174
 DB 52 AACGAGATGACGCTTACCAAGCCAGAACCAACGATGAGAGCTTGAAG 3

RESULT 12
 B0807302 640 bp mRNA linear EST 31-JUL-2002
 LOCUS B0807302
 DEFINITION NISC_K01010.y1 NCI CGAP Brn72 Macaca mulatta cDNA clone
 IMAGE:5330250 5', mRNA sequence.
 ACCESSION B0807302
 VERSION B0807302.1 GI:22031511
 SOURCE B0807302.1
 KEYWORDS EST.
 ORGANISM Macaca mulatta (rhesus monkey)
 Macaca mulatta
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea; Macaca.
 Cereopithecoidea; Macaca.
 NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
 Unpublished (1997)
 Contact: Robert Strausberg, Ph.D.
 Email: cga@bbs-rmail.nih.gov
 cDNA Library Preparation:
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium/LLNL
 DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC)
 Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: info@image.llnl.gov
 Plate: L14M1838 row: C column: 19
 Seq primer: M13Rpi reverse primer (AB1).
 Location/Qualifiers
 1..640
 /organism="Macaca mulatta"
 /mol_type="mRNA"
 /db_xref="taxon:9544"
 /clone="IMAGE:5330250"
 /cissue_type="hypothalamus"
 /lab_host="DH10B (phage-res;stanc)"
 /clone_lib="NCI CGAP Brn72"
 /note="Organ: brain; Vector: pCMV-SPORT6.1; Site 1: NotI; Site 2: EcoRV; Cloned unidirectionally. Primer: Qiagen dt. Average insert size 2.2 kb. Constructed by Invitrogen.
 Note: this is a NCI-CGAP Library."

ORIGIN
 Query Match 10.2%; Score 628.8; DB 5; Length 640;
 Best Local Similarity 98.9%; Pred. No. 1.4e-153;
 Matches 633; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 4073 GCCTCTGGCCCGCCATCCAGCCGAGAAAGAGTCTTCACTCCAGAGAAATTTAGTC 4132
 DB 1 GCCTCTGGCCCGCCATCCAGCCGAGAAAGAGTCTTCACTCCAGAGAAATTTAGTC 60
 QY 4133 GCGCTCTTAAGGAAGCATGACCAACATATTTCTTCAACGATTCAAGTGAAGTGAACA 4192
 DB 61 GCGCTCTTAAGGAAGCATGACCAACATATTTCTTCAACGATTCAAGTGAAGTGAACA 120

BG976452
 LOCUS BG976452 881 bp mRNA linear EST 12-JUN-2001
 DEFINITION 602846289F1 NCI_CGAP_Mam6 Mus musculus cDNA clone IMAGE:4976752 5',
 mRNA sequence.
 ACCESSION BG976452
 VERSION BG976452.1 GI:14364089
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 881)
 NIH-MGC <http://mgi.nci.nih.gov/>.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 CONTACT: Robert Strausberg, Ph.D.
 Email: cgabbs-remail.nih.gov
 Tissue Procurement: Jeffrey Green M.D.
 cDNA library Preparation: Life Technologies, Inc.
 cDNA library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
<http://image.llnl.gov>
 Place: LLNL0970 row: j column: 17
 High quality sequence stop: 841.
 Location/Qualifiers
 1..881
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="FVB/N"
 /db_xref="taxon:10090"
 /clone="IMAGE:4976752"
 /sex="female, virgin"
 /tissue_type="infiltrating ductal carcinoma"
 /dev_stage="5 months"
 /lab_host="DH10B"
 /clone_11b="NCI CGAP Mam6"
 /note="Organ: mammary; Vector: pCMV-SportE; Site 1: SalI;
 Site 2: NotI; Cloned unidirectionally. Primer: Oligo dt.
 library constructed by Life Technologies. Investigator
 providing samples: Jeffrey Green, M.D., NIH"

ORIGIN
 Query Match 9.8%; Score 603.4; DB 4; Length 881;
 Best Local Similarity 86.2%; Pred. No. 7.9e-147;
 Matches 726; Conservative 0; Mismatches 107; Indels 9; Gaps 5;

QY 5302 AATTACGATCTCTCATTTGGAACCAATAATTCTACGAATCGACATGAAGCAGTACAGC 5361
 DB 40 AAGAAACGATCTCTCATTTGGAACCAATAATTCTACGAATCGACATGAAGCAGTACAGC 99
 QY 5363 CTCGAGGATTCCTGATTAAGATGACCTTCTTGGCACCTGCTGTGTTGGCGCTCT 5421
 DB 100 CTTCGATGATTCCTGACAAAGACCACTTCTTGGCACCTGCTGTGTTGGCGCTCT 159
 QY 5422 TCCACAGCTTCCCTGCTCAATCGTGCAGTGAACAGCGCAGGCGAGGAGGAGTAC 5481
 DB 160 TCCACAGCTTCCCTGCTCAATCGTGCAGTGAACAGCGCAGGCGAGGAGGAGTAC 219
 QY 5482 TTGCTGTGTTTCCAGCAATTTGAGTGTTCGTGATTTCTTACGGAAGAGTACCGCACA 5541
 DB 220 CTGCTGTGTTTCCAGCAATTTGAGTGTTCGTGATTTCTTACGGAAGAGTACCGCACA 279
 QY 5542 GACGATTCATGAGTGGCTTACTTGGCTTGGCTTGGCTTACAGAGAACCTATCTGTT 5601
 DB 280 GATGATTCATGAGTGGCTTACTTGGCTTGGCTTGGCTTACAGAGAACCTATCTGTT 339
 QY 5602 GTGACCACTTCAACTCTGAGATTAATTGAGATCCAGACAGCTCTCAGAGGAGCC 5661
 DB 340 GTGACCACTTCAACTCTCCTGAGATTAATTGAGATCCAGACAGCTCTCAGAGGAGCC 399
 QY 5662 CCGTCCGAGCGTACCTGAGCATCCGGAACCGCGCTACTGAGGCGCTTCCATTTCTCA 5721

DB 400 CCGCCCGAGCATATCTGAAATTCGAAACCTTCGCTACTGGGCGCGGATTTCTCC 459
 QY 5722 GAGGATTTACTTGGCGTCTCTATACCAAGATTAATTAAGGATCAATTTGCTCAAGGA 5781
 DB 460 GAGGATTTACTTGGCGTCTCTATACCAAGATTAATTAAGGATCAATTTGCTCAAGGA 519
 QY 5782 AACCTGTGAAGAGTCCGCACTGAACAACAACCGGGGCGCGTCCACTTCGCGAGAGC 5841
 DB 520 AACCTGTGAAGAGTCAAGCACTGAGCAAGCAGCGGTCCTCCACTTCGCGAGAGC 579
 QY 5842 CCCAACAGGAGGCGCCACCACTGATCAACAGACATCAACCAAGGCGTGGCTCCAGC 5901
 DB 580 CCCAACAGGAGGCGCCCAACATCAACAGACATCAACCAAGGCGTGGCTCCAGC 638
 QY 5902 CCAGGCGCGCCGGAAGGCGCCAGCCAGCGCGGAGCCAGCAGCAGC-----CCACGCGT 5956
 DB 639 CCGGCGCCACCGGAAGGCGCCAGCCAGCGCGGAGCCAGCAGCAGCAGCAGCAGCAG 697
 QY 5957 ACCGCGAGGCGCGGACCGGAGCTGCGGAGGACCAAGTCTCTGCGCCCGCTGAGAGCAG 6016
 DB 698 ACAGAGAGGAGTGGACAGAGCTGCGGAGGACCAAGTCTCAAGCGCGCTTGAAGGCGG 757
 QY 6017 AGAAGTCCCCCGCGCGGATCTGAGCAGCAGGAGAGCGGTCCTCCCGCG-AGCTGTT 6074
 DB 758 AGAAGTCCCCCGCGCGGATCTGAGCAGGAGAGAGAGCGGTCCTCCAGGAGAGAGCTGTT 817
 QY 6075 TGAAGCAGCAGCAGGCGCGCGCTGCGGAGGCGGAGAGCGCGCTGCGCAGGT 6134
 DB 818 TGAAGCAGCAGCAGGCGCGCGCTGCGGAGGCGGAGAGCGCGCTGCGCAGGT 877
 QY 6135 GA 6136
 DB 878 AA 879

RESULT 15
 BP146990 698 bp mRNA linear EST 30-DEC-2003
 LOCUS BP146990 full-length enriched swine cDNA library, adult ovary Sus
 DEFINITION scrofa cDNA clone OVM10025B09 5', mRNA sequence.
 ACCESSION BP146990
 VERSION BP146990.1 GI:40396461
 KEYWORDS EST.
 SOURCE Sus scrofa (pig)
 ORGANISM Sus scrofa
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Cetartiodactyla; Suidae; Suidae; Sus.
 1 (bases 1 to 698)
 Uenishi, H., Eguchi, T., Suzuki, K., Sawazaki, T., Toki, D., Shinkai, H.,
 Okumura, N., Hamada, N. and Awata, T.
 PEDS (Pig EST Data Explorer): construction of a database for ESTs
 derived from porcine full-length cDNA libraries
 Nucleic Acids Res. 32 (1), D484-D488 (2004)
 CONTACT: Hirohide Uenishi
 Animal Genome Laboratory, Genome Research Department
 National Institute of Agrobiological Sciences
 2 Ikenodai, Tsukuba, Ibaraki 305-8602, Japan
 Tel: +81-29-838-8627
 Fax: +81-29-838-8627
 Email: huenishi@affrc.go.jp
 EST project with full-length enriched cDNA libraries carried out in
 Animal Genome Research Program (Japan) by National Institute of
 Agrobiological Sciences and STRF-Institute
 Single pass sequencing of clones derived from oligo-capped cDNA
 library
 Vector sequences were eliminated by RepeatMasker version 2002/07/13
 and crossmatch version 0.990319
 Low quality bases were trimmed based on the quality values.
 Location/Qualifiers
 1..698
 /organism="Sus scrofa"
 /mol_type="mRNA"

```

/db_xref="taxon:9823"
/clone="OVRM10025B09"
/tissue_type="ovary"
/dev_stage="adult"
/clone_lib="full-length enriched swine cdna library, adult
ovary"

```

ORIGIN

Query Match	9.8%	Score 602.6;	DB 5;	Length 698;
Best Local Similarity	91.4%;	Pred. No. 1.2e-146;		
Matches 638;	Conservative 0;	Mismatches 60;	Indels 0;	Gaps 0;

Qy	406	CAGAGACAGGTTTCATTTTTTGAGAAAGCGGAACTATTAATCTCGAAGACAAAGCCG	465
Db	1	CAGAGACAGGTTTCATTTTTTGAGAAAGACAAACACTTAATCCCGAGCAAGCCCT	60
Qy	466	TGATCCCCCAATTACAGTATGCCCTTTGAGACAAATAATCACTTTATCTGCTATGGAA	525
Db	61	TGATATCCCCCGAGTTTACAGTATGCCCTTTGAGACAAACATAACTTTATCTGCTATGGAA	120
Qy	526	TATACGCTGGAGGGGACTTGCTGTCACTTTTGAATGATATGAGGACACAGTTAGTGAA	585
Db	121	TATACGCTGGAGGGGATTTGCTGTCACTTTTGAATGATATGAGGACCAATTAGTGAA	180
Qy	586	AACCTGATACAGTTTAACTAGCTGAGCTGATTTTGGCTGTTCAACAGGTCATCTGATG	645
Db	181	AATATGATACAGTTTAACTAGCTGAATATGATTTTGGCTGTTCAACAGGTCATCAAGATG	240
Qy	646	GGATACGCTGATCGAGACATCAACCTGAGAAACATTCGCTGTGACCGGACAGGACACATC	705
Db	241	GGATATGATCAATCGAGACATCAAGCTGAGAACATTCATCAACGCAAGCGGACACATC	300
Qy	706	AAGCTGTGATTTTGGATCTGCCGCGAAATGAAATTCAAACAAGATGTAATGCCAA	765
Db	301	AAGCTGTGATCTTTGGATCAGCTGCTTAAGATGAACCTCAATTAAGATGTAATGCCAA	360
Qy	766	CTCCGATTTGGGACCCCGATTACATGGCTCTTGAAGTCTGACTGTGATGAACGGGAT	825
Db	361	CTCCGATTTGGGACCCCGATTACATGGCCCTTAAGCTGACTGTGATGAATGGGAT	420
Qy	826	GGAAAGGACACTACGCGCTGGACTGTGACTGTGATGAGGGCGTGAATGCCATAG	885
Db	421	GGAAAGAGTGCCTATGCGCTAGACTGTGATTTGGTGTGTCAAGTGGAGTTATTTGCTTAAG	480
Qy	886	ATGATTTATGAGGATGCCCTTGCGACAGGGAACTTCTGCCAAGACTTCAATPAACAT	945
Db	481	ATGATTTATGAGAGTCCCGCAATTCACGTAGGGAACTTCAAGCCAGAACTTCAATPAACATC	540
Qy	946	ATGAATTTCCAGCGGTTTTTGAATTTTCAATGACCCCAAGTGAGCAGTGACTTTCT	1005
Db	541	ATGAATTTCCAGCGGTTTTTGAAGTTTCAATGACCCCAAGTTAGCAGTGAATGCTT	600
Qy	1006	GATCTGATTCAAAGCTTTGTGTGGCGGCCAAGAGAGACTGAAGTTTGAAGGTCCTTTC	1065
Db	601	GATCTGATTCAAAGNTTATTTGTGTGGCCAGAAAGAGAGACTGAAGTTTGAAGGCTTTTC	660
Qy	1066	TGCCATCTTTCTCTCTPAAATATGACTGGAAACAACAT	1103
Db	661	TGTCATCTTTCTCTCTPAAATATGACTGGAAATAACAT	698

Search completed: May 16, 2005, 18:26:10
Job time : 12000 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 16, 2005, 02:34:29 ; Search time 1954 Seconds
(without alignments)
18677.178 Million cell updates/sec

Title: US-10-791-666-1
Perfect score: 6165
Sequence: 1 atgttgaggtcaacatcg.....gggaccagctctcagratataa 6165

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapept 1.0

Searched: 4390206 seqs, 2959870667 residues
Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
listing first 45 summaries

Database : N_Geneseq_16Dec04:.*
1: geneseq1980s:.*
2: geneseq1990s:.*
3: geneseq2000s:.*
4: geneseq2001as:.*
5: geneseq2001bs:.*
6: geneseq2002as:.*
7: geneseq2002bs:.*
8: geneseq2003as:.*
9: geneseq2003bs:.*
10: geneseq2003cs:.*
11: geneseq2003ds:.*
12: geneseq2004as:.*
13: geneseq2004bs:.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	6165	100.0	6165	6	ABQ78870	Abq78870 Human kin
2	6161.8	99.9	6298	6	AAD38864	Aad38864 Human kin
3	6160.2	99.9	6165	9	AAL55214	Aal55214 Human CRI
4	6160.2	99.9	8603	9	AAL55215	Aal55215 Human CRI
5	6144	99.7	8656	12	ADJ96544	Adj96544 Human cit
6	6142.2	99.6	6156	9	AAL55217	Aal55217 Human CRI
7	6126	99.4	6159	6	AAS06701	Aas06701 Polynucle
8	6121.8	99.3	6189	6	ABS63436	Abs63436 RHO/RAC-1
9	6121.8	99.3	6189	8	ADA05641	Ada05641 Human NOV
10	6120.2	99.3	6189	12	ADN62806	Adn62806 Human NOV
11	6105.4	99.0	6201	6	ABS63435	Abs63435 Human CDN
12	6105.4	99.0	6201	6	ADA05653	Ada05653 Human NOV
13	6091	98.8	6201	12	ADN62818	Adn62818 Human NOV
14	5837	94.7	5877	6	ABQ78871	Abq78871 Human kin
15	5661.4	91.8	6162	10	ADP60993	Adp60993 ORF of pa
16	5661.4	91.8	6574	6	AAD39191	Aad39191 Human MDP
17	5661.4	91.8	6574	10	ADP60992	Adp60992 Pain asro
18	5661.4	91.8	6574	13	ADQ89099	Adq89099 Human uro
19	3855	62.5	6609	3	AACT7568	Aac7568 Human ORF
20	2928.2	47.5	3131	4	ABA08361	Aba08361 Human RHO

21	2818.2	45.7	5261	9	AAL55216	Aal55216 Human CRI
22	2818.2	45.7	5261	12	ADM72200	Adm72200 Human TAS
23	2808.2	45.6	5251	10	ADD89966	Add89966 Human can
24	2682.8	43.5	2896	5	ABV30132	Abv30132 Human pro
25	2491.8	40.4	2693	10	ADE09823	Ade09823 Novel DNA
26	2476.6	40.2	2497	8	ADA05645	Ada05645 Human NOV
27	2421.6	39.3	2542	8	ADA05647	Ada05647 Human NOV
28	2379.4	38.6	2542	12	ADN62812	Adn62812 Human NOV
29	2329	37.8	2429	12	ADN62810	Adn62810 Human NOV
30	1847.8	30.0	1870	8	ADA05649	Ada05649 Human NOV
31	1847.4	30.0	1870	8	ADA05643	Ada05643 Human NOV
32	1844.2	29.9	1870	12	ADN62808	Adn62808 Human NOV
33	1843	29.9	1870	12	ADN62814	Adn62814 Human NOV
34	1791.2	29.1	1915	8	ADA05651	Ada05651 Human NOV
35	1766.6	28.7	1915	12	ADN62816	Adn62816 Human NOV
36	1401.4	22.7	1485	10	ABZ68725	Abz68725 Nucleotid
37	1401.4	22.7	1765	10	ABZ68726	Abz68726 Nucleotid
38	1399.8	22.7	2066	6	AAD26454	Aad26454 Human kin
39	1398.2	22.7	1515	8	ACA61394	Aca61394 CDNA enco
40	1398.2	22.7	1515	10	AAD59938	Aad59938 Human kin
41	1398.2	22.7	1515	12	ADO40591	Ado40591 Human kin
42	1088.8	17.7	2380	10	ABZ68776	Abz68776 Nucleotid
43	958.8	15.6	2162	8	ABX71191	Abx71191 Novel hum
44	957	15.5	957	11	ADN62729	Adn62729 Human cit
45	936.4	15.2	1048	5	AAS79753	Aas79753 DNA encod

ALIGNMENTS

RESULT 1	
ABQ78870	
ID	ABQ78870 standard; CDNA; 6165 BP.
XX	
AC	ABQ78870;
XX	
DT	10-OCT-2002 (first entry)
XX	
DE	Human kinase cDNA #1.
XX	
KW	Human, kinase; enzyme; serine-threonine kinase; nootropic; cytosolic;
KW	Citron rho-interacting kinase; gene therapy; mental disorder; cancer;
KW	gene; ss.
XX	
OS	Homo sapiens.
XX	
FT	Key
FT	Location/Qualifiers
FT	1..6165
FT	/*tag= a
FT	/product= "kinase"
FT	/replace (5218..G)
FT	/*tag= b
FT	/strand name= "Single nucleotide polymorphism"
FT	variation
FT	/*tag= b
FT	/strand name= "Single nucleotide polymorphism"
FT	variation
FT	/*tag= b
FT	/strand name= "Single nucleotide polymorphism"
XX	
PD	WO200259325-A2.
XX	
PA	01-AUG-2002.
XX	
XX	(LEXI-) LEXICON GENETICS INC.
XX	
PI	Yu X, Miranda M, Fridde CJ;
XX	
DR	WPI: 2002-599796/64.
DR	P-PDB; ABB81927.
XX	
PT	Novel polynucleotide encoding human proteins that are structurally

PT similar to animal kinases, useful for drug screening, diagnosis, in gene
PT therapy of disorders and diseases e.g. cancer and pharmacogenomic
PT applications.

PS Claim 1; Page 37-39; 50pp; English.

XX The invention relates to a novel human protein that shares structural
CC similarity with animal kinases, including serine-threonine kinases,
CC particularly Cdkron rho-interacting kinases. The proteins of the
CC invention have mitogenic and cytoskeletal activity. The polypeptides may
CC have a use in gene therapy. The encoded novel polypeptides are useful for
CC generating antibodies, as reagents in diagnostic assays, for identifying
CC other cellular gene products related to NHP and as reagents in assays for
CC screening for compounds that are useful in the treatment of mental,
CC biological or medical disorders and diseases including cancer. The
CC sequence encodes a novel human kinase of the invention

XX Sequence 6165 BP; 1735 A; 1550 C; 1679 G; 1201 T; 0 U; 0 Other;

Query Match 100.0%; Score 6165; DB 6; Length 6165;
Best Local Similarity 100.0%; Pred. No. 0;

Matches 6165; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 ATGTTGAAGTTCAATATGAGCGGGAATCTTGGATGCTGGTGGTGAACCAT 60
Db 1 ATGTTGAAGTTCAATATGAGCGGGAATCTTGGATGCTGGTGGTGAACCAT 60
QY 61 GCCAGCCGGGCTTCAGAGCTGAATCTGTTCTTCCAGGGGAAACCACTTATGACTCA 120
Db 61 GCCAGCCGGGCTTCAGAGCTGAATCTGTTCTTCCAGGGGAAACCACTTATGACTCA 120
QY 121 CAGAGATGTCCTCTCTTCCCGAAGGATATTAAGTCCCTTTGTTCTTTGAA 180
Db 121 CAGAGATGTCCTCTCTTCCCGAAGGATATTAAGTCCCTTTGTTCTTTGAA 180
QY 181 GAATGACGTAGCTCTGATGAAGATTAAAGACGTGAACAATTGTCGGAATAT 240
Db 181 GAATGACGTAGCTCTGATGAAGATTAAAGACGTGAACAATTGTCGGAATAT 240
QY 241 TCCGACACCAATGCTGATTAACAGAGCTCAAGCTTCGCAAGAGCTTGAAGTGA 300
Db 241 TCCGACACCAATGCTGATTAACAGAGCTCAAGCTTCGCAAGAGCTTGAAGTGA 300
QY 301 AGCTCTGATGAGTGTGCTCTTGTGCAAGTGAAGTGAAGAGAAAGCAACCGGG 360
Db 301 AGCTCTGATGAGTGTGCTCTTGTGCAAGTGAAGTGAAGAGAAAGCAACCGGG 360
QY 361 GACATCTATGTAAGATGAAGAGAGAGCTTATTTGCCAGAGAGAGGTTTCA 420
Db 361 GACATCTATGTAAGATGAAGAGAGAGCTTATTTGCCAGAGAGAGGTTTCA 420
QY 421 TTTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 480
Db 421 TTTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 480
QY 481 CAGTATGCTTTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 540
Db 481 CAGTATGCTTTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 540
QY 541 GACTTGTCTGCACTTTGAATGATGAAGAGAGAGAGAGAGAGAGAGAGAGAG 600
Db 541 GACTTGTCTGCACTTTGAATGATGAAGAGAGAGAGAGAGAGAGAGAGAGAG 600
QY 601 TACCTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 660
Db 601 TACCTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 660
QY 661 GACATCAAGCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 720
Db 661 GACATCAAGCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 720
QY 721 GGAATCTGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 780
Db 721 GGAATCTGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 780
```

```
Db 721 GGAATCTGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 780
QY 781 CCAGATTATACATGCTCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 840
Db 781 CCAGATTATACATGCTCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 840
QY 841 GGCTTGAGCTGTGATCTGTGTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 900
Db 841 GGCTTGAGCTGTGATCTGTGTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 900
QY 901 TCCCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 960
Db 901 TCCCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 960
QY 961 TTTTGAAGATTTTCCAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1020
Db 961 TTTTGAAGATTTTCCAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1020
QY 1021 TTTTGTGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1080
Db 1021 TTTTGTGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1080
QY 1081 TCTTAAATTTGATCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1140
Db 1081 TCTTAAATTTGATCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1140
QY 1141 TCTGAGATGACACCTTCAATTTTGAATGAACAGAGAGAGAGAGAGAGAGAGAG 1200
Db 1141 TCTGAGATGACACCTTCAATTTTGAATGAACAGAGAGAGAGAGAGAGAGAGAG 1200
QY 1201 CCGTGCAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1260
Db 1201 CCGTGCAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1260
QY 1261 TACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1320
Db 1261 TACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1320
QY 1321 CCGTGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1380
Db 1321 CCGTGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1380
QY 1381 TCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1440
Db 1381 TCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1440
QY 1441 GAGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1500
Db 1441 GAGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1500
QY 1501 TCCCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1560
Db 1501 TCCCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1560
QY 1561 TTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1620
Db 1561 TTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1620
QY 1621 GATATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1680
Db 1621 GATATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1680
QY 1681 GTGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1740
Db 1681 GTGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1740
QY 1741 CCGAGTGTATCTTCAAGATCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1800
Db 1741 CCGAGTGTATCTTCAAGATCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1800
QY 1801 CCGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1860
Db 1801 CCGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1860
```

QY 1861 GTGGAGAAATATGCGAAACTGGAGAGATCAATGCTGAGCAGCAGCTCAAAATTACAGAG 1920
DB 1861 GTGGAGAAATATGCGAAACTGGAGAGATCAATGCTGAGCAGCAGCTCAAAATTACAGAG 1920
QY 1921 CTCGAAGAGAACTGGAGAGGCTGTATAAGCCAGCAGCAGAGGCGACCGAGCTGCTGAG 1980
DB 1921 CTCGAAGAGAACTGGAGAGGCTGTATAAGCCAGCAGCAGAGGCGACCGAGCTGCTGAG 1980
QY 1981 AATATCCGCGAGCCAAAGAGACGAGCCGAGAGGAGCTGGAGAACTGCAAGAACCGAGAG 2040
DB 1981 AATATCCGCGAGCCAAAGAGACGAGCCGAGAGGAGCTGGAGAACTGCAAGAACCGAGAG 2040
QY 2041 GATTCTTCTGAAAGCATCAGAAAAGACTGAGAGCTGAGAACTGAGAACTGAGAACTGAG 2100
DB 2041 GATTCTTCTGAAAGCATCAGAAAAGACTGAGAGCTGAGAACTGAGAACTGAGAACTGAG 2100
QY 2101 GAGAAACAGGTAAGAGACTGAGAGCAGTGAAGGCTGAGAGAAACAATGAGAGATGAC 2160
DB 2101 GAGAAACAGGTAAGAGACTGAGAGCAGTGAAGGCTGAGAGAAACAATGAGAGATGAC 2160
QY 2161 ATCCAGACAAATATCCCAACAGATCAGCAGATGGCTGATAAATTCTGAGCTGAGAGAG 2220
DB 2161 ATCCAGACAAATATCCCAACAGATCAGCAGATGGCTGATAAATTCTGAGCTGAGAGAG 2220
QY 2221 AAAATCGGAGAGGCCCAAGTCTCAGCCAGCAGCCTAGAGTGCACCTGAAAACAGAAAGAG 2280
DB 2221 AAAATCGGAGAGGCCCAAGTCTCAGCCAGCAGCCTAGAGTGCACCTGAAAACAGAAAGAG 2280
QY 2281 CAGCACTATGAGAGAAAGATTAAAGTGTGAGCAATGATTAAGAAAGAACTGCTGAGAC 2340
DB 2281 CAGCACTATGAGAGAAAGATTAAAGTGTGAGCAATGATTAAGAAAGAACTGCTGAGAC 2340
QY 2341 AAGGAGACATGAGAGCAATGATGAGAGACAGAGAGAGAGGCCCAATGAGAAAGAGAG 2400
DB 2341 AAGGAGACATGAGAGCAATGATGAGAGACAGAGAGAGAGGCCCAATGAGAAAGAGAG 2400
QY 2401 ATTCTCAGCGAACAAGAGGCGATGATCAATGCTATGATTCAGAGATCAAGATCCCTGAA 2460
DB 2401 ATTCTCAGCGAACAAGAGGCGATGATCAATGCTATGATTCAGAGATCAAGATCCCTGAA 2460
QY 2461 CAGAGATTTGTGAGACTGTCTGAAAGCCAAATTAATCTGAGAGAAATAGCAGCTTTTAA 2520
DB 2461 CAGAGATTTGTGAGACTGTCTGAAAGCCAAATTAATCTGAGAGAAATAGCAGCTTTTAA 2520
QY 2521 CAAAGGAACTGAAGGCCCAAGAGAGATGATTTCTGAATCAGAGCAACGAATTTTAC 2580
DB 2521 CAAAGGAACTGAAGGCCCAAGAGAGATGATTTCTGAATCAGAGCAACGAATTTTAC 2580
QY 2581 CTGGAGACAGAGCTGGAGAGTTGAGAGCCCAAGAACTGGAGAGAGCAGCTGAGAG 2640
DB 2581 CTGGAGACAGAGCTGGAGAGTTGAGAGCCCAAGAACTGGAGAGAGCAGCTGAGAG 2640
QY 2641 AAGATCAGCCACCAAGACCAAGTACAGAAATGAGTCTGAGAACTGGAGACAAAGATTG 2700
DB 2641 AAGATCAGCCACCAAGACCAAGTACAGAAATGAGTCTGAGAACTGGAGACAAAGATTG 2700
QY 2701 CGGAGAGTCAAGTCTGAGAGCAGAGAGCAAGAACTGAGAGCTGAGAGCTGAGAGAG 2760
DB 2701 CGGAGAGTCAAGTCTGAGAGCAGAGAGCAAGAACTGAGAGCTGAGAGCTGAGAGAG 2760
QY 2761 CTACAGCTCTCCCTGAGAGAGCGAGAGTCAAGTGTGAGAGCTGAGAGCTGAGAGAG 2820
DB 2761 CTACAGCTCTCCCTGAGAGAGCGAGAGTCAAGTGTGAGAGCTGAGAGCTGAGAGAG 2820
QY 2821 GCCCTGAGAGAGCGAGTCTGAGAGCGAGAGCAAGAGCTGAGAGAGCAAGAGAGAGT 2880
DB 2821 GCCCTGAGAGAGCGAGTCTGAGAGCGAGAGCAAGAGCTGAGAGAGCAAGAGAGAGT 2880
QY 2881 GAAAGAGAGATCCAGGACATCAAGGCAATGAGATGAAATCCAGAGCGAAATTTGAGTCT 2940
DB 2881 GAAAGAGAGATCCAGGACATCAAGGCAATGAGATGAAATCCAGAGCGAAATTTGAGTCT 2940

QY 2941 CTTGCTTAACAGCTGTACTGTAAATCAGAGACTGAGAGAGAGCTTAAACAGCTGAGAG 3000
DB 2941 CTTGCTTAACAGCTGTACTGTAAATCAGAGACTGAGAGAGAGCTTAAACAGCTGAGAG 3000
QY 3001 GACAAAGCTGAATCTCAACCAACCAAACTTACTTGTCCAAACCAATCGATGAGAGCTTCT 3060
DB 3001 GACAAAGCTGAATCTCAACCAACCAAACTTACTTGTCCAAACCAATCGATGAGAGCTTCT 3060
QY 3061 GGGGCCAACGACGAGATTTGTACAATGCGAAGTGAAGTGAACCACTTCGCGCGGAGATC 3120
DB 3061 GGGGCCAACGACGAGATTTGTACAATGCGAAGTGAAGTGAACCACTTCGCGCGGAGATC 3120
QY 3121 ACCGAACGAGAGATGACGCTTACAGCCAGAGCAACCAATGAGAGCTTGAAGACCAAG 3180
DB 3121 ACCGAACGAGAGATGACGCTTACAGCCAGAGCAACCAATGAGAGCTTGAAGACCAAG 3180
QY 3181 TGCACCATGCTGAGAGGAAACAGGTCAATGGAATTTGAGAGGCTTAAACGATGAGCTGAGA 3240
DB 3181 TGCACCATGCTGAGAGGAAACAGGTCAATGGAATTTGAGAGGCTTAAACGATGAGCTGAGA 3240
QY 3241 AAGAGCGGAGAGTGGAGAGGCTGAGAGAGCGTCTGAGGATGAGAAATCCAGTTTGA 3300
DB 3241 AAGAGCGGAGAGTGGAGAGGCTGAGAGAGCGTCTGAGGATGAGAAATCCAGTTTGA 3300
QY 3301 TGTGAGTTGAGAGCTGAGAGAGATGCTGGAACAACGAGAAACAGAGCGGCGAGAGCC 3360
DB 3301 TGTGAGTTGAGAGCTGAGAGAGATGCTGGAACAACGAGAAACAGAGCGGCGAGAGCC 3360
QY 3361 GATCAGCGGATCAACCGAGTCTCCGCAAGTGTGAGAGTGGAGAGGAGAGCAAGAGCT 3420
DB 3361 GATCAGCGGATCAACCGAGTCTCCGCAAGTGTGAGAGTGGAGAGGAGAGCAAGAGCT 3420
QY 3421 GAGATTCGCTCTGAGCAGAGCTCTCAAGAGAGAGAGCTGAGAGCGGAGAGCTCTCT 3480
DB 3421 GAGATTCGCTCTGAGCAGAGCTCTCAAGAGAGAGAGCTGAGAGCGGAGAGCTCTCT 3480
QY 3481 GACAGCTCAATGACTGAGAGAAAGAGAGATGCTTGTGAATGATGCCCAAGCTTA 3540
DB 3481 GACAGCTCAATGACTGAGAGAAAGAGAGATGCTTGTGAATGATGCCCAAGCTTA 3540
QY 3541 CAGAGAGCTGAGAGACTGAACGAGAGCTCAACAGAGGCTTCTGAGAGAGCAAGCCAA 3600
DB 3541 CAGAGAGCTGAGAGACTGAACGAGAGCTCAACAGAGGCTTCTGAGAGAGCAAGCCAA 3600
QY 3601 TTACAGCAGAGATGAGACTGAGAGAAATCAGATTTTCCGTCTGACTCAAGAGCTGAA 3660
DB 3601 TTACAGCAGAGATGAGACTGAGAGAAATCAGATTTTCCGTCTGACTCAAGAGCTGAA 3660
QY 3661 GAAAGCTTAGATCGGCTGATCTTCTGAAGACAGAAAGAACTGAGAGATCACTG 3720
DB 3661 GAAAGCTTAGATCGGCTGATCTTCTGAAGACAGAAAGAACTGAGAGATCACTG 3720
QY 3721 GAAACCTTCAAGTCTCTATTTCTGATGAAAGTGAATTTGAGAGCACTAATTTCTCA 3780
DB 3721 GAAACCTTCAAGTCTCTATTTCTGATGAAAGTGAATTTGAGAGCACTAATTTCTCA 3780
QY 3781 CAAACCAACTCAATGATTTTCTGCAAGCCAAATGAGACCACTGCTTAAAGAAAG 3840
DB 3781 CAAACCAACTCAATGATTTTCTGCAAGCCAAATGAGACCACTGCTTAAAGAAAG 3840
QY 3841 GTTCTCTGAGAGTACAAGAGCTGAGAGCTGAGAGAGAGAGAGAGAGTCTGCTGCA 3900
DB 3841 GTTCTCTGAGAGTACAAGAGCTGAGAGCTGAGAGAGAGAGAGAGAGTCTGCTGCA 3900
QY 3901 GAGTTAAGAGAGAGCTTCAAGAGAGAGAGAGAGAGTCTGAGAGAGAGAGAGAGT 3960
DB 3901 GAGTTAAGAGAGAGCTTCAAGAGAGAGAGAGAGAGTCTGAGAGAGAGAGAGAGT 3960
QY 3961 GCCCAGCGAAAG 4020
DB 3961 GCCCAGCGAAAG 4020
QY 4021 ATGCGCATGTCGCGCATGTCGCGTGCACAGAGCAGAGCCAGTGCATGAGAGCTGCTG 4080


```
Db 4021 ATGGCCATGTCGCGATCGGATCGCCAGAGCAACGAGCCAGGCGCATAGGCTGTG 4080
Qy 4081 GCCCGCCATCCAGCCGAGAAAGAGTCTTCAACTCCAGAGAAATTTAGTCGGGCTTT 4140
Db 4081 GCCCGCCATCCAGCCGAGAAAGAGTCTTCAACTCCAGAGAAATTTAGTCGGGCTTT 4140
Qy 4141 AAGGAAGCGATGCAACCAATATTCCTCAACGATTCAACGTAGACATGAACTCGAGCC 4200
Db 4141 AAGGAAGCGATGCAACCAATATTCCTCAACGATTCAACGTAGACATGAACTCGAGCC 4200
Qy 4201 ACAAGGTGCTGTGTCTGATGATACCGTGCATCTTGGAGCGGAGATCCAAATGCTC 4260
Db 4201 ACAAGGTGCTGTGTCTGATGATACCGTGCATCTTGGAGCGGAGATCCAAATGCTC 4260
Qy 4261 GAATGTGAGTGTGTCAACCCAGTGTCCAGTGTGCGAGCCACTGCGGCTTG 4320
Db 4261 GAATGTGAGTGTGTCAACCCAGTGTCCAGTGTGCGAGCCACTGCGGCTTG 4320
Qy 4321 CCTGTGAAATATGCGACACATTCACCGAGGCTTCTGCGGTGACAAATGAACTCCCA 4380
Db 4321 CCTGTGAAATATGCGACACATTCACCGAGGCTTCTGCGGTGACAAATGAACTCCCA 4380
Qy 4381 GGTCTCAGACCAAGAGCCGAGCGCTTGACCTGGAAGGTGATGAAGTGTGCC 4440
Db 4381 GGTCTCAGACCAAGAGCCGAGCGCTTGACCTGGAAGGTGATGAAGTGTGCC 4440
Qy 4441 AGGAATTAACAAACGAGGACGAAAGGCTGGAGACGAAATGATATGCTCGAGGATCA 4500
Db 4441 AGGAATTAACAAACGAGGACGAAAGGCTGGAGACGAAATGATATGCTCGAGGATCA 4500
Qy 4501 AAAGTCTCTATTTATGACATGAAGCCAGAAAGCTGACAGAGCGCGGTGAAATTT 4560
Db 4501 AAAGTCTCTATTTATGACATGAAGCCAGAAAGCTGACAGAGCGCGGTGAAATTT 4560
Qy 4561 GAGCTGTGCTTCCGACGCGGAGATGATCATTCATGTGCGGTGTGTTCCGAATC 4620
Db 4561 GAGCTGTGCTTCCGACGCGGAGATGATCATTCATGTGCGGTGTGTTCCGAATC 4620
Qy 4621 GCAAAATACAGCCAAAGCAGATGTCCCATCATACTAGTAAGATGAAATCTCACCCGACAC 4680
Db 4621 GCAAAATACAGCCAAAGCAGATGTCCCATCATACTAGTAAGATGAAATCTCACCCGACAC 4680
Qy 4681 ACCGTGTGCGCCGCGGAGAAACCTCTATCTTGTAGCTTCCAGCTTCCCTGACAAACGCC 4740
Db 4681 ACCGTGTGCGCCGCGGAGAAACCTCTATCTTGTAGCTTCCAGCTTCCCTGACAAACGCC 4740
Qy 4741 TGGGTACGCGCTTAGAATCAGTGTGCGAGGTGGAGAGTTCTAGGGAAAAAGAGAA 4800
Db 4741 TGGGTACGCGCTTAGAATCAGTGTGCGAGGTGGAGAGTTCTAGGGAAAAAGAGAA 4800
Qy 4801 GCTGATGCTAACTGCTTGGAAAATCTCCGCTGAAACTGGAAGGTATGACGCTTAGAC 4860
Db 4801 GCTGATGCTAACTGCTTGGAAAATCTCCGCTGAAACTGGAAGGTATGACGCTTAGAC 4860
Qy 4861 ATGAATGCAACGCTGCTTCACTGATGACAGAGTGTGTGTGTGGGCAACGAGAAAGGCTC 4920
Db 4861 ATGAATGCAACGCTGCTTCACTGATGACAGAGTGTGTGTGTGGGCAACGAGAAAGGCTC 4920
Qy 4921 TAGCGCCGTAATGCTTGAATAAATCTCCCTAACCCAGTGTCCAGGAATTTGAGACAGTCTTC 4980
Db 4921 TAGCGCCGTAATGCTTGAATAAATCTCCCTAACCCAGTGTCCAGGAATTTGAGACAGTCTTC 4980
Qy 4981 CAAATTTATATTAATCAAGACCTGAGAGCTACTATGATAGCAGAGAAAGCGGACA 5040
Db 4981 CAAATTTATATTAATCAAGACCTGAGAGCTACTATGATAGCAGAGAAAGCGGACA 5040
Qy 5041 CTGTGTCTTGTGAGAGTGAAGAAAGTGAAGTGTGCTGCGCAAGTCCCACTGTGCTGC 5100
Db 5041 CTGTGTCTTGTGAGAGTGAAGAAAGTGAAGTGTGCTGCGCAAGTCCCACTGTGCTGC 5100
Qy 5101 CAGCCCGCATCTCACCAATTTTGAAGCTGTCAAGGCTGCACTTGTGTGGGGCA 5160
Db 5101 CAGCCCGCATCTCACCAATTTTGAAGCTGTCAAGGCTGCACTTGTGTGGGGCA 5160

Db 5101 CAGCCCGCATCTCACCAATTTTGAAGCTGTCAAGGCTGCACTTGTGTGGGGCA 5160
Qy 5161 GGCAGAAATGAGAAAGGGGCTCTGCATCTGTGACAGCCATGCCCAGCAAGTGTGATCTC 5220
Db 5161 GGCAGAAATGAGAAAGGGGCTCTGCATCTGTGACAGCCATGCCCAGCAAGTGTGATCTC 5220
Qy 5221 CGCTTACAGAAAACCTCAGCAAAATCTGCAATCCGGAAGAGATGAGACTTCAGAGCCC 5280
Db 5221 CGCTTACAGAAAACCTCAGCAAAATCTGCAATCCGGAAGAGATGAGACTTCAGAGCCC 5280
Qy 5281 TGCAGCTGTATCACTTCAACCAATTAACAGATCTCATTTGGAAACCAATAATCTACGAA 5340
Db 5281 TGCAGCTGTATCACTTCAACCAATTAACAGATCTCATTTGGAAACCAATAATCTACGAA 5340
Qy 5341 ATGACATGAAAGAGATACAGCTCGAGGAATTCCTGATTAAGATGACATTCCTTGGCA 5400
Db 5341 ATGACATGAAAGAGATACAGCTCGAGGAATTCCTGATTAAGATGACATTCCTTGGCA 5400
Qy 5401 CTGTGTGTGTTGCGGCTCTTCCAAACAGCTTCCGTGTCAATCGAGAGTGAACAGC 5460
Db 5401 CTGTGTGTGTTGCGGCTCTTCCAAACAGCTTCCGTGTCAATCGAGAGTGAACAGC 5460
Qy 5461 GCAGGGCAGCGAGAGAGATCTTGTGTGTTCCAGAAATTTGAGATGTTCTGGAATCT 5520
Db 5461 GCAGGGCAGCGAGAGAGATCTTGTGTGTTCCAGAAATTTGAGATGTTCTGGAATCT 5520
Qy 5521 TACGGAAGAGTGTAGCCGACAGACGATCTCAAGTGTGAGTCTTACCTTGTGGCTTGGC 5580
Db 5521 TACGGAAGAGTGTAGCCGACAGACGATCTCAAGTGTGAGTCTTACCTTGTGGCTTGGC 5580
Qy 5581 TACAGAAACCTATCTGTTGTGACCACTTCAACTCACTGAAATTAATGATTCAG 5640
Db 5581 TACAGAAACCTATCTGTTGTGACCACTTCAACTCACTGAAATTAATGATTCAG 5640
Qy 5641 GCAAGCTCTCTCAGCAGGAGACCTTGTGCGAGAGTACTGGAATCCCGAGCTGAC 5700
Db 5641 GCAAGCTCTCTCAGCAGGAGACCTTGTGCGAGAGTACTGGAATCCCGAGCTGAC 5700
Qy 5701 CTGGGCTGCTGCAATTCCTCAGAGAGATTTACTTGGCTCTCATACAGATTAATTA 5760
Db 5701 CTGGGCTGCTGCAATTCCTCAGAGAGATTTACTTGGCTCTCATACAGATTAATTA 5760
Qy 5761 AAGGTATTTGCTGCAAGGAAAACCTGTGAAGAGTCCGGGCACTGGAACCAACCGGGGC 5820
Db 5761 AAGGTATTTGCTGCAAGGAAAACCTGTGAAGAGTCCGGGCACTGGAACCAACCGGGGC 5820
Qy 5821 CGGTCCACCTCCCGAGAGAGCCCAACAGCGAGCCCAACGATACAGAGACATC 5880
Db 5821 CGGTCCACCTCCCGAGAGAGCCCAACAGCGAGCCCAACGATACAGAGACATC 5880
Qy 5881 ACCAAGCGGTGAGCTTCCAGCCACGCGCCGAGAGCCGAGCAACCGCGAGAGCCA 5940
Db 5881 ACCAAGCGGTGAGCTTCCAGCCACGCGCCGAGAGCCGAGCAACCGCGAGAGCCA 5940
Qy 5941 AGCACAACCCACCGCTTACCGCGAGGAGCGGACCGAGCTGCGAGGAGCAAGTCTCTGGC 6000
Db 5941 AGCACAACCCACCGCTTACCGCGAGGAGCGGACCGAGCTGCGAGGAGCAAGTCTCTGGC 6000
Qy 6001 CGCCCCCTGAGAGCGAGAAAGTCCCGCGCGGATACACGACACCGAGAGAGAGGTCC 6060
Db 6001 CGCCCCCTGAGAGCGAGAAAGTCCCGCGCGGATACACGACACCGAGAGAGAGGTCC 6060
Qy 6061 CCCGCGAGGCTGTGTAAGAGACGACGAGGAGCGGCTGCGCTGCGGAGCGGTGAGACC 6120
Db 6061 CCCGCGAGGCTGTGTAAGAGACGACGAGGAGCGGCTGCGCTGCGGAGCGGTGAGACC 6120
Qy 6121 CCGCTGTCCAGGTGAAACAGGTCTGGAGCAAGTCTTCAAGTAA 6165
Db 6121 CCGCTGTCCAGGTGAAACAGGTCTGGAGCAAGTCTTCAAGTAA 6165
Qy 6165 CCGCTGTCCAGGTGAAACAGGTCTGGAGCAAGTCTTCAAGTAA 6165
Db 6165 CCGCTGTCCAGGTGAAACAGGTCTGGAGCAAGTCTTCAAGTAA 6165
```

RESULT 2
AAD38864

QY 961 TTTTGAATTTCCAGATGAGCCCAAGTGAAGTCACTTTCTTGATCTGATTCAAAGC 1020
Db 1015 TTTTGAATTTCCAGATGAGCCCAAGTGAAGTCACTTTCTTGATCTGATTCAAAGC 1074
QY 1021 TTGTTGCGGCGCAGAAAGAGAGACTGAAGTTTGAAGGCTTTTGCTGCGCATCTTTCTTC 1080
Db 1075 TTGTTGCGGCGCAGAAAGAGAGACTGAAGTTTGAAGGCTTTTGCTGCGCATCTTTCTTC 1134
QY 1081 TCTAAATTTGACTGGAACAACATTTCTGTAAGTCTCTCCCTCTTCGTTCCACCTCAAG 1140
Db 1135 TCTAAATTTGACTGGAACAACATTTCTGTAAGTCTCTCCCTCTTCGTTCCACCTCAAG 1194
QY 1141 TCTGAAGATGACACTCTCAATTTTGTGAAGCAAGAAAGAAATTTGTGGGTTTCACTCT 1200
Db 1195 TCTGAAGATGACACTCTCAATTTTGTGAAGCAAGAAAGAAATTTGTGGGTTTCACTCT 1254
QY 1201 CCGTGCAAGTGAAGCCCTCAAGGCTTCTCGGTTGAAGAACTGCGGTTTGTGGGGTTTTCG 1260
Db 1255 CCGTGCAAGTGAAGCCCTCAAGGCTTCTCGGTTGAAGAACTGCGGTTTGTGGGGTTTTCG 1314
QY 1261 TACAGCAAGGCACTGGGGAATTTGTAGATCTGAGTCTGTGTGTGCGGTTCTGSACTCC 1320
Db 1315 TACAGCAAGGCACTGGGGAATTTGTAGATCTGAGTCTGTGTGTGCGGTTCTGSACTCC 1374
QY 1321 CCGGCCAAGACTAGCTTCATGSAAGAAAGAACTTCTCATCAAAAGCAAGAGCTACAGAC 1380
Db 1375 CCGGCCAAGACTAGCTTCATGSAAGAAAGAACTTCTCATCAAAAGCAAGAGCTACAGAC 1434
QY 1381 TCTCAGGACAAGTGTCAAGATGAGAGAGAAATGACCCGGTTTACATCGGAGAGTGTCA 1440
Db 1435 TCTCAGGACAAGTGTCAAGATGAGAGAGAAATGACCCGGTTTACATCGGAGAGTGTCA 1494
QY 1441 GAGGTGAGGCTGTGCTTGTAGTCAAGAGAGTGAAGGCTTCTGAGACTCAGAGA 1500
Db 1495 GAGGTGAGGCTGTGCTTGTAGTCAAGAGAGTGAAGGCTTCTGAGACTCAGAGA 1554
QY 1501 TCCCTCTGGAAGAGACTTTGCTTACTTCAATCAAGAAATGCAAGTACTTAAAGCGAAGT 1560
Db 1555 TCCCTCTGGAAGAGACTTTGCTTACTTCAATCAAGAAATGCAAGTACTTAAAGCGAAGT 1614
QY 1561 TTGAGCAAGCAAGATGAGAGTGTCCAGAGAGATGACAAAGCACTGACGTTTCCAT 1620
Db 1615 TTGAGCAAGCAAGATGAGAGTGTCCAGAGAGATGACAAAGCACTGACGTTTCCAT 1674
QY 1621 GATATCAGAGAGAGAGCCCGAAGCTTCAAGAAATCAAGAGCAGAGTACAGGCTCA 1680
Db 1675 GATATCAGAGAGAGAGCCCGAAGCTTCAAGAAATCAAGAGCAGAGTACAGGCTCA 1734
QY 1681 GTGGAAGAAATGAGTTGATGATCAATCAAGTGAAGAGATTTTGTCTCAGCAAGAGA 1740
Db 1735 GTGGAAGAAATGAGTTGATGATCAATCAAGTGAAGAGATTTTGTCTCAGCAAGAGA 1794
QY 1741 CCGAGTGAATCTTCAAGATCTGAGTGAAGAGTCTCGGCTTGTCTGTAAGAAATTCAG 1800
Db 1795 CCGAGTGAATCTTCAAGATCTGAGTGAAGAGTCTCGGCTTGTCTGTAAGAAATTCAG 1854
QY 1801 CCGAAAGCAGCAAGATGTCAAGTAACTGTTGAAGGCTTAAGGATCAAGGAGAGCTGAA 1860
Db 1855 CCGAAAGCAGCAAGATGTCAAGTAACTGTTGAAGGCTTAAGGATCAAGGAGAGCTGAA 1914
QY 1861 GTGGGAGAAATGCGCAAACTGAGAGATCAATGCTGAGACAGCTCAAAAATTCAGAG 1920
Db 1915 GTGGGAGAAATGCGCAAACTGAGAGATCAATGCTGAGACAGCTCAAAAATTCAGAG 1974
QY 1921 CTCCAGAGAAACTGAGAGAGGCTGTAAAGCCAGACCGAGGSCCAACGAGCTGTCAG 1980
Db 1975 CTCCAGAGAAACTGAGAGAGGCTGTAAAGCCAGACCGAGGSCCAACGAGCTGTCAG 2034
QY 1981 AATATCCGCAAGAGAGAGGAGCCGAGAGGAGCTGGAAGAGCTTGCAAGACGAGAG 2040
Db 2035 AATATCCGCAAGAGAGAGGAGCCGAGAGGAGCTGGAAGAGCTTGCAAGACGAGAG 2094

QY 2041 GATTTCTTGAAGGATCAGAAAGAGAGCTGTGGAAGCTGAGAGACCCCGCATTTCTG 2100
Db 2095 GATTTCTTGAAGGATCAGAAAGAGAGCTGTGGAAGCTGAGAGACCCCGCATTTCTG 2154
QY 2101 GAGAACAGGTTAAAGAGACTAGAGACCATGAGAGGTGAGAGAAACAGACTGAAGATGAC 2160
Db 2155 GAGAACAGGTTAAAGAGACTAGAGACCATGAGAGGTGAGAGAAACAGACTGAAGATGAC 2214
QY 2161 ATCCAGCAAAATCCCAACAGATCCAGCAGATGCTGATTAATTTCTGAGCTGAAAG 2220
Db 2215 ATCCAGCAAAATCCCAACAGATCCAGCAGATGCTGATTAATTTCTGAGCTGAAAG 2274
QY 2221 AAACATCGGAGGCCCCAAGTCTCAGCCAGACCTTGAAGTGAACCTGAAACGAAAGAG 2280
Db 2275 AAACATCGGAGGCCCCAAGTCTCAGCCAGACCTTGAAGTGAACCTGAAACGAAAGAG 2334
QY 2281 CAGCACTATGAGAAATTAAGTTAAGTTGGAACAATCAGATTAAGAAAGACCTGGCTGAC 2340
Db 2335 CAGCACTATGAGAAATTAAGTTAAGTTGGAACAATCAGATTAAGAAAGACCTGGCTGAC 2394
QY 2341 AAGGAGACACTGAGAGAACTGATGACAGAGCAAGAGAGGAGCCCATGAGAGGCAAA 2400
Db 2395 AAGGAGACACTGAGAGAACTGATGACAGAGCAAGAGAGGAGCCCATGAGAGGCAAA 2454
QY 2401 ATTCTAGCGAACAAGAGGCGATGATCAATGCTATGGAATTCAGATTCAGATCCCTGAA 2460
Db 2455 ATTCTAGCGAACAAGAGGCGATGATCAATGCTATGGAATTCAGATTCAGATCCCTGAA 2514
QY 2461 CAGAGATTGTGGAATGTCTGAAGCCAAATTAACCTGACAGCAAAATGACGCTTTTACC 2520
Db 2515 CAGAGATTGTGGAATGTCTGAAGCCAAATTAACCTGACAGCAAAATGACGCTTTTACC 2574
QY 2521 CAAAGAACATGAGAGGCCCAAGAAAGATGATTTTCTGAATCTCAGGCAACAGAAATTTAC 2580
Db 2575 CAAAGAACATGAGAGGCCCAAGAAAGATGATTTTCTGAATCTCAGGCAACAGAAATTTAC 2634
QY 2581 CTGGAACAACAGGCTGGGAAGTTGGAGGCCCAAGAACCGAAACTGAGAGGACGCTGAG 2640
Db 2635 CTGGAACAACAGGCTGGGAAGTTGGAGGCCCAAGAACCGAAACTGAGAGGACGCTGAG 2694
QY 2641 AAGATCAGCACCAAGACCCACAGTGAACAAGATTCGGCTGTGGAATCTGAGACCAAGATTG 2700
Db 2695 AAGATCAGCACCAAGACCCACAGTGAACAAGATTCGGCTGTGGAATCTGAGACCAAGATTG 2754
QY 2701 CCGGAGTCACTCTGAGACAGAGAGCAAGAACTGGAATCAAGGCCAGCTCACAAGAG 2760
Db 2755 CCGGAGTCACTCTGAGACAGAGAGCAAGAACTGGAATCAAGGCCAGCTCACAAGAG 2814
QY 2761 CTACAGCTTCCCTGCAAGAGCGCAAGTCAAGTGAACGCTTGAAGGCTGACGAGCG 2820
Db 2815 CTACAGCTTCCCTGCAAGAGCGCAAGTGAACGCTTGAAGGCTGACGAGCG 2874
QY 2821 GCCCTGAGAGAGCAGCTTGCCAGAGGGAAGACAGAGCTGGAAGAGCCACACAGAGCT 2880
Db 2875 GCCCTGAGAGAGCAGCTTGCCAGAGGGAAGACAGAGCTGGAAGAGCCACACAGAGCT 2934
QY 2881 GAAAGAGAGATCCAGGCACTCAGGCACTAGAGATGAAGTCAAGGCAAAATTTGATGCT 2940
Db 2935 GAAAGAGAGATCCAGGCACTCAGGCACTAGAGATGAAGTCAAGGCAAAATTTGATGCT 2994
QY 2941 CTTGCTAACAGCTGTATCTGTATCAACAGCTGAGAGAGCAGCTTAAACCAAGCTGACCGAG 3000
Db 2995 CTTGCTAACAGCTGTATCTGTATCAACAGCTGAGAGAGCAGCTTAAACCAAGCTGACCGAG 3054
QY 3001 GACAACGCTGAATCAACCAAACTTTACTGTTCCAAACCACTGATGAGGCTTCT 3060
Db 3055 GACAACGCTGAATCAACCAAACTTTACTGTTCCAAACCACTGATGAGGCTTCT 3114
QY 3061 GCGGCCAAAGCAGAGATTGTACAATCTGGAAGTGAAGTGAACATCTCCGCGGAGAGATC 3120
Db 3115 GCGGCCAAAGCAGAGATTGTACAATCTGGAAGTGAAGTGAACATCTCCGCGGAGAGATC 3174
QY 3121 ACGGAACAGAGATGCAAGCTTACAGCAAGCAAGAAAGATGAGAGCTTGAAGACCAAG 3180

Db 3175 ACGGACGAGATGACGCTTACCAAGCAGAGAAACGATGAGGCTTGAAGACACG 3234
Qy 3181 TGCACATGCTGAGAGAAACAGCTATGATTTGGAGCCCTAAACGATGAGCTGTAGAA 3240
Db 3235 TGACCAATGCTGAGAGAAACAGCTATGATTTGGAGCCCTAAACGATGAGCTGTAGAA 3294
Qy 3241 AAAGAGCGGCACTGGAGAGGCTGAGAGAGCTCTGGGTGATGAGAAATCCAGTTTGA 3300
Db 3295 AAAGAGCGGCACTGGAGAGGCTGAGAGAGCTCTGGGTGATGAGAAATCCAGTTTGA 3354
Qy 3301 TGTGGGTTGAGAGCTGCAAGAGATGCTGGACACGAGAAACGAGACAGGGCCGAGAGCC 3360
Db 3355 TGTGGGTTGAGAGCTGCAAGAGATGCTGGACACGAGAAACGAGACAGGGCCGAGAGCC 3414
Qy 3361 GATCAGCGGATCACCGAGTCTCCGACAGTGGTGGAGCTGGCACTGAAAGACACAGGCT 3420
Db 3415 GATCAGCGGATCACCGAGTCTCCGACAGTGGTGGAGCTGGCACTGAAAGACACAGGCT 3474
Qy 3421 GAGATTCTGCTGTGACAGAGCTCTCAAGAGACAGAAAGCTGAAGGCTGAGAGCTCTCT 3480
Db 3475 GAGATTCTGCTGTGACAGAGCTCTCAAGAGACAGAAAGCTGAAGGCTGAGAGCTCTCT 3534
Qy 3481 GACAACTCAATGACCTGGAGAGAAAGCATGCTATGCTTGAATGAATGCCCCAGCTTA 3540
Db 3535 GACAACTCAATGACCTGGAGAGAAAGCATGCTATGCTTGAATGAATGCCCCAGCTTA 3594
Qy 3541 CAGCAGAACTGGAGCTGAAACGAGAGCTCAACAGAGGCTTCTGAAAGACAGCCAA 3600
Db 3595 CAGCAGAACTGGAGCTGAAACGAGAGCTCAACAGAGGCTTCTGAAAGACAGCCAA 3654
Qy 3601 TTACAGCAGCAGATGACCTGCAAGAAATCAATTTTCGCTGACTCAAGCATGCA 3660
Db 3655 TTACAGCAGCAGATGACCTGCAAGAAATCAATTTTCGCTGACTCAAGCATGCA 3714
Qy 3661 GAAGCTTAAGTCCGGCTGATCTTACTGAACACAGAAAGATGACTTGGAGTACAGCTG 3720
Db 3715 GAAGCTTAAGTCCGGCTGATCTTACTGAACACAGAAAGATGACTTGGAGTACAGCTG 3774
Qy 3721 GAAGAACTGAGTCTCTATCTCATGAAAGGTGAAATGAAAGGCACTATTCTCA 3780
Db 3775 GAAGAACTGAGTCTCTATCTCATGAAAGGTGAAATGAAAGGCACTATTCTCA 3834
Qy 3781 CAACCAAACTCATTTGATTTTCTGCAAGCCAAATGACCAACTGCTAAAGAAAG 3840
Db 3835 CAACCAAACTCATTTGATTTTCTGCAAGCCAAATGACCAACTGCTAAAGAAAG 3894
Qy 3841 GTTCTCTGCACTAATGAGCTGAGCTGAGCTGAGAGAGAAAGTCTGCTGTGCA 3900
Db 3895 GTTCTCTGCACTAATGAGCTGAGCTGAGCTGAGAGAGAAAGTCTGCTGTGCA 3954
Qy 3901 GAGCTAAGAGAAAGCTTCTGAGAAACCCGCACTGAGCTCCGGCCGGAGAGAGCT 3960
Db 3955 GAGCTAAGAGAAAGCTTCTGAGAAACCCGCACTGAGCTCCGGCCGGAGAGAGCT 4014
Qy 3961 GCCCAGCGAAAGACGAGACCAACCACTCATCAGCAGGACCGGAGGACAG 4020
Db 4015 GCCCAGCGAAAGACGAGACCAACCACTCATCAGCAGGACCGGAGGACAG 4074
Qy 4021 ATGCGCATGTCGGCATGTCGGGTGCGCAGAGACCAAGCCAGTGCATGAGCTGTG 4080
Db 4075 ATGCGCATGTCGGCATGTCGGGTGCGCAGAGACCAAGCCAGTGCATGAGCTGTG 4134
Qy 4081 GCCCGGCACTCAGCCGAGAAAGAGATCTTCACTCAGAGAAATTAATGTCGGCTTT 4140
Db 4135 GCCCGGCACTCAGCCGAGAAAGAGATCTTCACTCAGAGAAATTAATGTCGGCTTT 4194
Qy 4141 AAGGAAAGCATGACCACAATATTCCTCAACGATTTCAAGTGAAGCTGAACATCGAGCC 4200
Db 4195 AAGGAAAGCATGACCACAATATTCCTCAACGATTTCAAGTGAAGCTGAACATCGAGCC 4254
Qy 4201 ACAAAGT 4260

Db 4255 ACAAAGT 4314
Qy 4261 GAATGTGAGTATGT 4320
Db 4315 GAATGTGAGTATGT 4374
Qy 4321 CCGTGAATATGT 4380
Db 4375 CCGTGAATATGT 4434
Qy 4381 GGTCTCAGACCAAGAGCCCAAGAGCTTGTGACCTGTGAAAGGTGTGATGAAAGTGTCCC 4440
Db 4435 GGTCTCAGACCAAGAGCCCAAGAGCTTGTGACCTGTGAAAGGTGTGATGAAAGTGTCCC 4494
Qy 4441 AAGATATCAAAACGAGACGCAAGCTGGAGACAGAAATGATCAATGCTCTGAGAGATCA 4500
Db 4495 AAGATATCAAAACGAGACGCAAGCTGGAGACAGAAATGATCAATGCTCTGAGAGATCA 4554
Qy 4501 AAAGTCTCTATTTATGACATGAAAGCTGGAAGAGCTGGAGAGAGCCGGTGGAGAAATTT 4560
Db 4555 AAAGTCTCTATTTATGACATGAAAGCTGGAAGAGCTGGAGAGAGCCGGTGGAGAAATTT 4614
Qy 4561 GAGCTGCTTCCGACGAGGAGTGTATTAATTCATGATGCTGCGGTGGTGTCTCCGAACTC 4620
Db 4615 GAGCTGCTTCCGACGAGGAGTGTATTAATTCATGATGCTGCGGTGGTGTCTCCGAACTC 4674
Qy 4621 GCAAATACAGCAAGACAGATGTCCCATATCATATGAAATGGAATCTCACCCGACACC 4680
Db 4675 GCAAATACAGCAAGACAGATGTCCCATATCATATGAAATGGAATCTCACCCGACACC 4734
Qy 4681 ACTGCTGGCCCGGAGAACCTCTTACTTCTGTACTGCTCCAGCTTCTGACAAACAGCCG 4740
Db 4735 ACTGCTGGCCCGGAGAACCTCTTACTTCTGTACTGCTCCAGCTTCTGACAAACAGCCG 4794
Qy 4741 TGGGTCAACGCTTGAATGATGTTGTGAGGTGGAGAGTTTCTAGAGAAAGACAGAA 4800
Db 4795 TGGGTCAACGCTTGAATGATGTTGTGAGGTGGAGAGTTTCTAGAGAAAGACAGAA 4854
Qy 4801 GCTGATGCTAATGCTTGTGAAATCTCCCTGCTGAACTGGAAGGTATGACCGTCTAGAC 4860
Db 4855 GCTGATGCTAATGCTTGTGAAATCTCCCTGCTGAACTGGAAGGTATGACCGTCTAGAC 4914
Qy 4861 ATGAACTGCAACGCTGCCCTTCACTGACAGGTGTGTGTGGGACACGAGAAAGGCTC 4920
Db 4915 ATGAACTGCAACGCTGCCCTTCACTGACAGGTGTGTGTGGGACACGAGAAAGGCTC 4974
Qy 4921 TAGCCCTGAATGCTTGAAGAAATCTCCCTAACCATGCTCCAGGAAATGAGAGCTTTC 4980
Db 4975 TAGCCCTGAATGCTTGAAGAAATCTCCCTAACCATGCTCCAGGAAATGAGAGCTTTC 5034
Qy 4981 CAATTTATTTATCAAGAGACTGAGAACTCTCATGATGACAGAGAAAGAGCGGCA 5040
Db 5035 CAATTTATTTATCAAGAGACTGAGAACTCTCATGATGACAGAGAAAGAGCGGCA 5094
Qy 5041 CTGTGCTTGTGAGCTGAAAGAAAGTGAACAGTCCCTGAGCCCATCTGCTGCTC 5100
Db 5095 CTGTGCTTGTGAGCTGAAAGAAAGTGAACAGTCCCTGAGCCCATCTGCTGCTC 5154
Qy 5101 CAGCCGCACTTCAACCAATTTTGAAGCTGTCAAGGCTGCCACTTGTTTGGGGCA 5160
Db 5155 CAGCCGCACTTCAACCAATTTTGAAGCTGTCAAGGCTGCCACTTGTTTGGGGCA 5214
Qy 5161 GGGCAAGTTGAGAAAGGAGCTCTGATGTGTGACCAATGCCCAGCAAAAGTGTCTCTC 5220
Db 5215 GGGCAAGTTGAGAAAGGAGCTCTGATGTGTGACCAATGCCCAGCAAAAGTGTCTCTC 5274
Qy 5221 CGCTAACAGAAACCTCAGCAAAATCTGATCTCGAAGAGATGAGACTCAGAGCCC 5280
Db 5275 CGCTAACAGAAACCTCAGCAAAATCTGATCTCGAAGAGATGAGACTCAGAGCCC 5334
Qy 5281 TGAGAGCTGTATCACTTCAACCAATTAAGATATCTCTTATGGAACCAATTAATTTAGAA 5340
Db 5335 TGAGAGCTGTATCACTTCAACCAATTAAGATATCTCTTATGGAACCAATTAATTTAGAA 5394

```

QY 5341 ATGACATGAAGAGTACAGCTGAGAGAAATTCCTGATAGATGACATTCTTGGCA 5400
DB 5395 ATGACATGAAGAGTACAGCTGAGAGAAATTCCTGATAGATGACATTCTTGGCA 5454
QY 5401 CTTGCTGTGTTTGGCGCTCTTCAAGAGCTTCCCTGCTCAATCGAGAGTGAACAGC 5460
DB 5455 CTTGCTGTGTTTGGCGCTCTTCAAGAGCTTCCCTGCTCAATCGAGAGTGAACAGC 5514
QY 5461 GCAGGGCAGAGAGAGTACTTGCTGTGTTTCAAGAAATTTGGAGTGTTCGTGAATCT 5520
DB 5515 GCAGGGCAGAGAGAGTACTTGCTGTGTTTCAAGAAATTTGGAGTGTTCGTGAATCT 5574
QY 5521 TAGGAAAGAGTACCGGACAGAGATCTGAAGAGTGGCTTACTTTGGCTTTGGCC 5580
DB 5575 TAGGAAAGAGTACCGGACAGAGATCTGAAGAGTGGCTTACTTTGGCTTTGGCC 5634
QY 5581 TAGAGAAACCTATCTGTTTGTGACCACTTCAACTCACTCGAAGATTTGAGATCCAG 5640
DB 5635 TAGAGAAACCTATCTGTTTGTGACCACTTCAACTCACTCGAAGATTTGAGATCCAG 5694
QY 5641 GCACGCTCTTCAAGAGGAGCCCTGCGCGAGCTGACCTGACATCCCGAGTAC 5700
DB 5695 GCACGCTCTTCAAGAGGAGCCCTGCGCGAGCTGACCTGACATCCCGAGTAC 5754
QY 5701 CTGGGCGCTGCGCAATTTCTCTGAGAGGATTTACTTGCGCTCTCTATACAGATTAATTA 5760
DB 5755 CTGGGCGCTGCGCAATTTCTCTGAGAGGATTTACTTGCGCTCTCTATACAGATTAATTA 5814
QY 5761 AGGGTCAATTTGTGCAAGAGGAAACCTGCTGAAGAGTCCGCGACTGAAACCAACCGGGAGC 5820
DB 5815 AGGGTCAATTTGTGCAAGAGGAAACCTGCTGAAGAGTCCGCGACTGAAACCAACCGGGAGC 5874
QY 5821 CCGTCCACCTCCGCGAGAGGAGCCCAACAGAGGAGCCCAACGTAACAAGAGACATC 5880
DB 5875 CCGTCCACCTCCGCGAGAGGAGCCCAACAGAGGAGCCCAACGTAACAAGAGACATC 5934
QY 5881 ACCAAGCGGTGTCCTTCCAGCCCAAGGCGCGCCGAGAGGCCCAACCCGCGAGAGCCA 5940
DB 5935 ACCAAGCGGTGTCCTTCCAGCCCAAGGCGCGCCGAGAGGCCCAACCCGCGAGAGCCA 5994
QY 5941 AGCAGACCCCAACCGCTACCGCGAGAGGAGGAGCCGAGCTGCGAGAGCAAGTCTCTGCGC 6000
DB 5995 AGCAGACCCCAACCGCTACCGCGAGAGGAGGAGCCGAGCTGCGAGAGCAAGTCTCTGCGC 6054
QY 6001 CCGCCCTCTGAGAGGAGAGTCCCGCGCGAGTACTAGACAGCGCGAGAGCGGCTCC 6060
DB 6055 CCGCCCTCTGAGAGGAGAGTCCCGCGCGAGTACTAGACAGCGCGAGAGCGGCTCC 6114
QY 6061 CCGCGAGAGGCTGTTTGAAGACAGACAGAGGAGCGCGCTGCTGCGGAGAGCCGTGAGACC 6120
DB 6115 CCGCGAGAGGCTGTTTGAAGACAGACAGAGGAGCGCGCTGCTGCGGAGAGCCGTGAGACC 6174
QY 6121 CCGCTGTCCAGGTTGAACAGAGTCTGAGGACCACTTTCAGTATTA 6165
DB 6175 CCGCTGTCCAGGTTGAACAGAGTCTGAGGACCACTTTCAGTATTA 6219

```

```

KW obesity; comorbidities; cancer; anorexia; cachexia; bulimia;
KM central nervous system disorder; chronic obstructive pulmonary disease;
XX diabetes; pain; gene; ds.
OS Homo sapiens.
FH Key Location/Qualifiers
FT CDS 1..6165
FT /product= "Human CRK protein"
PN WC2003004523-A1.
XX
PD 16-JAN-2003.
XX
PF 28-JUN-2002; 2002MO-EP007156.
XX
PR 02-JUL-2001; 2001US-0301841P.
PR 11-DEC-2001; 2001US-0338651P.
PR 25-APR-2002; 2002US-0375014P.
XX
PA (FARB ) BAYER AG.
XX
PI Zhu Z;
XX
XX WPI; 2003-221576/21.
DR P-PSDB; A026959.
XX
PT New human citron rho/rac-interacting kinase (CRK) polypeptide and
PT polynucleotide, useful in preventing, ameliorating or treating diseases
PT associated with human CRK dysfunction, e.g. obesity, diabetes or
PT Alzheimer's disease.
XX
XX Example 1; Fig 1; 237bp; English.
XX
XX The invention relates to an isolated polynucleotide encoding a human
XX citron rho/rac-interacting kinase polypeptide. The isolated
XX polynucleotide comprises a 6165 or 8603 base pair sequence, given in the
XX specification. The human citron rho/rac-interacting kinase (CRK)
XX polypeptide and polynucleotide are useful in preventing, ameliorating, or
XX treating diseases associated with human CRK dysfunction such as obesity
XX and obesity-associated comorbidities (e.g. hypertension, coronary artery
XX disease, hyperlipidaemia, stroke, gout, osteoarthritis, some types of
XX cancer including endometrial, breast, prostate and colon cancer),
XX anorexia, cachexia, bulimia, central nervous system disorders (e.g. mood
XX disorders, anxiety disorder, Parkinson's disease or Alzheimer's
XX disease), chronic obstructive pulmonary disease, or diabetes. These can
XX also be used to treat pain associated with the disorders. The human CRK
XX polypeptide is also useful in diagnostic assays or in genetic testing.
XX The expression vector or the reagent is useful in preparing a medicament
XX for modulating the activity of a human CRK in a disease, e.g. obesity, a
XX central nervous system disorder, or chronic obstructive pulmonary
XX disease. The fusion protein is useful for generating antibodies against a
XX CRK polypeptide and for use in various assay systems. The methods are
XX useful in producing and detecting the polynucleotide and polypeptide and
XX in screening for agents that modulate the activity of the human CRK
XX polypeptide. This polynucleotide sequence represents a DNA sequence
XX encoding a human CRK protein of the invention

```

Sequence 6165 BP; 1735 A; 1549 C; 1680 G; 1201 T; 0 U; 0 Other;

Query Match 99.9%; Score 6160.2; DB 9; Length 6165;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 6162; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

QY 1 ATGTTGAAGTTCAATATGAGACCGCGAATCTTTGATGCTGCTGCTGCAACCAATT 60
DB 1 ATGTTGAAGTTCAATATGAGACCGCGAATCTTTGATGCTGCTGCTGCAACCAATT 60
QY 61 GCACGCGGGGCTCAGGCTGAATGTTCTTCCAGGGGAAACCAACCTTATGACTCAA 120
DB 61 GCACGCGGGGCTCAGGCTGAATGTTCTTCCAGGGGAAACCAACCTTATGACTCAA 120

```


Dh 2281 CAGCACTATGAGAAAAGATTAAAGTGTGGACAATCAGATTAAGAAAGACCTGGCTGAC 2340
Qy 2341 AAGGAGACATGGGAACATGATCGAAGACACAGAGAGAGGCCCATGAGAGGGCCAAA 2400
Db 2341 AAGGAGACATGGGAACATGATCGAAGACACAGAGAGAGGCCCATGAGAAAGGCGAAA 2400
Qy 2401 ATTCTCAGCGAAGCAGAGGCGATGATCAATGCTATGATTCAGAGATCAGATCCCTGAAA 2460
Db 2401 ATTCTCAGCGAAGCAGAGGCGATGATCAATGCTATGATTCAGAGATCAGATCCCTGAAA 2460
Qy 2461 CAGAGATTGTGAACTGTCTGAAGCCAAATMACTTCGAGCAATATAGCAGCTTTTACC 2520
Db 2461 CAGAGATTGTGAACTGTCTGAAGCCAAATMACTTCGAGCAATATAGCAGCTTTTACC 2520
Qy 2521 CAAAAGAACTGAAGGCCCAAGAAAGATGATTTCTGAAGCTCAGGCAACAGAAATTTTAC 2580
Db 2521 CAAAAGAACTGAAGGCCCAAGAAAGATGATTTCTGAAGCTCAGGCAACAGAAATTTTAC 2580
Qy 2581 CTGGAGACACAGGCTGGGAAGTTGGAGGCCAGAACTGGAGAGAGAGCTGGAG 2640
Db 2581 CTGGAGACACAGGCTGGGAAGTTGGAGGCCAGAACTGGAGAGAGAGCTGGAG 2640
Qy 2641 AAGATCGCCACCAAGACACACAGTGAAGAAATCGCTGCTGGAATCGAGACAAGATTG 2700
Db 2641 AAGATCGCCACCAAGACACACAGTGAAGAAATCGCTGCTGGAATCGAGACAAGATTG 2700
Qy 2701 CGGAGAGTCACTCTAGAGACAGAGAGCAGAAACTGGAGCTCAAGCCGACCTCAAGAG 2760
Db 2701 CGGAGAGTCACTCTAGAGACAGAGAGCAGAAACTGGAGCTCAAGCCGACCTCAAGAG 2760
Qy 2761 CTACAGCTCTCCCTGACAGAGCGCGAGTCAAGTTGACACCTTCGACGGCTGACCGGCG 2820
Db 2761 CTACAGCTCTCCCTGACAGAGCGCGAGTCAAGTTGACACCTTCGACGGCTGACCGGCG 2820
Qy 2821 GCCCTGAGAGAGCAGCTTTCGCGAGGCGAGACAGAGCTGGAGAGACACAGCAGAACT 2880
Db 2821 GCCCTGAGAGAGCAGCTTTCGCGAGGCGAGACAGAGCTGGAGAGAGACACAGCAGAACT 2880
Qy 2881 GAAAGAGAGATCCAGGCACTCAGCGCACATGAGATGAATTCAGCGCAAAATTTGATGCT 2940
Db 2881 GAAAGAGAGATCCAGGCACTCAGCGCACATGAGATGAATTCAGCGCAAAATTTGATGCT 2940
Qy 2941 CTTGCTGAACAGCTGTACTGTATCACAAGCTTGGAGAGAGATTAACCGACCTGACCGAG 3000
Db 2941 CTTGCTGAACAGCTGTACTGTATCACAAGCTTGGAGAGAGATTAACCGACCTGACCGAG 3000
Qy 3001 GACAAAGCTGAATCTCAACAAACAAACCTTCTACTTCTCCAAACAATCGATGAGGCTTCT 3060
Db 3001 GACAAAGCTGAATCTCAACAAACAAACCTTCTACTTCTCCAAACAATCGATGAGGCTTCT 3060
Qy 3061 GGGCGCCACGAGAGATTGTAACTGCGAAGTGAAGTGAACCATCTCCGCGGGAGATC 3120
Db 3061 GGGCGCCACGAGAGATTGTAACTGCGAAGTGAAGTGAACCATCTCCGCGGGAGATC 3120
Qy 3121 ACGGAAACGAGATGAGAGCTTACAGCCAGAAAGCAATGAGAGCTTGAAGACCAAG 3180
Db 3121 ACGGAAACGAGATGAGAGCTTACAGCCAGAAAGCAATGAGAGCTTGAAGACCAAG 3180
Qy 3181 TGCACCATCTGAGAGAAACAGGTCAATGATTTGAGAGGCTTAAACGATGAGGCTAGAA 3240
Db 3181 TGCACCATCTGAGAGAAACAGGTCAATGATTTGAGAGGCTTAAACGATGAGGCTAGAA 3240
Qy 3241 AAAAGAGCGGAGTGGAGGCTTGGAGAGCGTCTTGGAGTGAAGAAATCCAGTTTGAAG 3300
Db 3241 AAAAGAGCGGAGTGGAGGCTTGGAGAGCGTCTTGGAGTGAAGAAATCCAGTTTGAAG 3300
Qy 3301 TGTCCGGGTTTCAGAGGCTGAGAGAGATGCTGACACCGAAGAAACAGAGCGGCGAGAGCC 3360
Db 3301 TGTCCGGGTTTCAGAGGCTGAGAGAGATGCTGACACCGAAGAAACAGAGCGGCGAGAGCC 3360
Qy 3361 GATCAGCGGATCAGGAGTCTGGCCAGAGTGTGGAGCTGGCAGTGAAGAGGACAAAGGCT 3420
Db 3361 GATCAGCGGATCAGGAGTCTGGCCAGAGTGTGGAGCTGGCAGTGAAGAGGACAAAGGCT 3420

Qy 3421 GAGATTCTGCTTTCGACAGGCTTTCAAAGACAGAAAGCTGAAGGCCGAGAGCTCTCT 3480
Db 3421 GAGATTCTGCTTTCGACAGGCTTTCAAAGACAGAAAGCTGAAGGCCGAGAGCTCTCT 3480
Qy 3481 GACAGCTCATGACCTGGAGAGAAAGCAATGCTATGCTTGAATGAAATGCCCCAGCTTA 3540
Db 3481 GACAGCTCATGACCTGGAGAGAAAGCAATGCTATGCTTGAATGAAATGCCCCAGCTTA 3540
Qy 3541 CAGCAGAGCTGGAGACTGAAACGAGAGCTCAACAGAGGCTTTCGAAAGACCAAGCCAAA 3600
Db 3541 CAGCAGAGCTGGAGACTGAAACGAGAGCTCAACAGAGGCTTTCGAAAGACCAAGCCAAA 3600
Qy 3601 TTACAGAGCAGATGGAAGCTTCGAGAAATACATTTTCCGTCTGACTCAGAGCTGCAA 3660
Db 3601 TTACAGAGCAGATGGAAGCTTCGAGAAATACATTTTCCGTCTGACTCAGAGCTGCAA 3660
Qy 3661 GAAGCTCTGAGTGGGCTCATCTCAAGAGACAGAAAGAGTGAAGTGAAGTGAAGTGAAG 3720
Db 3661 GAAGCTCTGAGTGGGCTCATCTCAAGAGACAGAAAGAGTGAAGTGAAGTGAAGTGAAG 3720
Qy 3721 GAAACATTCAAGTCTCTATTCTCATGAAAGGTGAAGAGGCACTATTCTCAA 3780
Db 3721 GAAACATTCAAGTCTCTATTCTCATGAAAGGTGAAGAGGCACTATTCTCAA 3780
Qy 3781 CAAACCAAACTCATGATTTTTCGAAAGCCAAATGAGCAACCTGCTAAGAAAGAAAG 3840
Db 3781 CAAACCAAACTCATGATTTTTCGAAAGCCAAATGAGCAACCTGCTAAGAAAGAAAG 3840
Qy 3841 GTTCCCTGAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 3900
Db 3841 GTTCCCTGAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 3900
Qy 3901 GAGCTAGAGAAAGCCCTTCAGAAAGCCCGCATGAGCTCCGATCCGCGGAGAGAGCT 3960
Db 3901 GAGCTAGAGAAAGCCCTTCAGAAAGCCCGCATGAGCTCCGATCCGCGGAGAGAGCT 3960
Qy 3961 GCCCACCAGAAAGCAACGAGCAACCCACACCATTCACCGCAGCCACCGGAGGACAGAG 4020
Db 3961 GCCCACCAGAAAGCAACGAGCAACCCACACCATTCACCGCAGCCACCGGAGGACAGAG 4020
Qy 4021 ATGCGCATGCTCCGCAATCGTGGGTGCGCAGAGACACAGCCAGTGCATGAGCCCTGAG 4080
Db 4021 ATGCGCATGCTCCGCAATCGTGGGTGCGCAGAGACACAGCCAGTGCATGAGCCCTGAG 4080
Qy 4081 GCCCGCCATCCAGCCGAGAAAGAGTCTTCACTCCAGAGAAATTTAGTGGGCTTCT 4140
Db 4081 GCCCGCCATCCAGCCGAGAAAGAGTCTTCACTCCAGAGAAATTTAGTGGGCTTCT 4140
Qy 4141 AAGGAAACGATGACCAATATTCCTCACCGATTCAACGTAGAGCTGAACATGCGAGCC 4200
Db 4141 AAGGAAACGATGACCAATATTCCTCACCGATTCAACGTAGAGCTGAACATGCGAGCC 4200
Qy 4201 ACAAAGT 4260
Db 4201 ACAAAGT 4260
Qy 4261 GAATGTCAGGTATGT 4320
Db 4261 GAATGTCAGGTATGT 4320
Qy 4321 CCTGCTGAATATGCAACAATTCTTCCAGAGGCTTTCGCGGTGACAAATGAATCTCCCA 4380
Db 4321 CCTGCTGAATATGCAACAATTCTTCCAGAGGCTTTCGCGGTGACAAATGAATCTCCCA 4380
Qy 4381 GGTCTTCAGACCAAGAGCCCAAGAGCTTTCGAGAGGCTTTCGAGAGGCTTTCGAGAGG 4440
Db 4381 GGTCTTCAGACCAAGAGCCCAAGAGCTTTCGAGAGGCTTTCGAGAGGCTTTCGAGAGG 4440
Qy 4441 AGGAATTAACAAACGAGGACAGCAAGGCTGGGACAGAGAGTACATTTGCTGGAGGGATCA 4500
Db 4441 AGGAATTAACAAACGAGGACAGCAAGGCTGGGACAGAGAGTACATTTGCTGGAGGGATCA 4500

QY 4501 AAAGTCTCATTTTATGACAAATGAGCCAGAGAACTGACAGAGAGCCCGTGGAAATTT 4560
 Db 4501 AAATCTCTATTATGACAAATGAGCCAGAGAACTGACAGAGAGCCCGTGGAAATTT 4560
 QY 4561 GAGCTGTGCTTCCCGAGCGGGGATGTATCTATTATGATGCTGCTTGGTGTCTTCCGAATC 4620
 Db 4561 GAGCTGTGCTTCCCGAGCGGGGATGTATCTATTATGATGCTGCTTGGTGTCTTCCGAATC 4620
 QY 4621 GCAAAATACAGCAAGCAAGATGTCCATACATCTAGATGAATGAATCTCACCCGACACC 4680
 Db 4621 GCAAAATACAGCAAGCAAGATGTCCATACATCTAGATGAATGAATCTCACCCGACACC 4680
 QY 4681 ACCGTCTGCGCCCGGAGAAACCTCTACTGCTAGCTCCAGCTTCCCTGACAAACAGGCG 4740
 Db 4681 ACCGTCTGCGCCCGGAGAAACCTCTACTGCTAGCTCCAGCTTCCCTGACAAACAGGCG 4740
 QY 4741 TGGGTCAACCGCTTATGATGATGCTGTGAGAGTGGAGATTTCTAGGGAAGAAAGAGAA 4800
 Db 4741 TGGGTCAACCGCTTATGATGATGCTGTGAGAGTGGAGATTTCTAGGGAAGAAAGAGAA 4800
 QY 4801 GCTGATCTAACTGCTTGGAAACTCCCTGCTGAACCTGAAAGGTGATGACCTGCTAGAC 4860
 Db 4801 GCTGATCTAACTGCTTGGAAACTCCCTGCTGAACCTGAAAGGTGATGACCTGCTAGAC 4860
 QY 4861 ATGAATCTGACGCTGCTTCACTGACAGAGTGTGTTGTGGGCAACCGAGAGAGGCTC 4920
 Db 4861 ATGAATCTGACGCTGCTTCACTGACAGAGTGTGTTGTGGGCAACCGAGAGAGGCTC 4920
 QY 4921 TACGCCCTGAAATGCTTGGAAAACTCCCTTAACCATGCTCCAGAAATTGGAGAGCTTC 4980
 Db 4921 TACGCCCTGAAATGCTTGGAAAACTCCCTTAACCATGCTCCAGAAATTGGAGAGCTTC 4980
 QY 4981 CAAATTTATATTTATCAAGAGCCTGAGAGAGCTACTCATGATAGAGAGAGAGAGCGGACA 5040
 Db 4981 CAAATTTATATTTATCAAGAGCCTGAGAGAGCTACTCATGATAGAGAGAGAGAGCGGACA 5040
 QY 5041 CTGTGTCTTGTGACGTGAAGAAAGTGAACAGTCCCTGCGCCAGTCCACCTGCTGCGC 5100
 Db 5041 CTGTGTCTTGTGACGTGAAGAAAGTGAACAGTCCCTGCGCCAGTCCACCTGCTGCGC 5100
 QY 5101 CAGGCCGACATCTGACCAATTTTGAAGCTGTCAGAGGCTGCCACTGTTTGGGACA 5160
 Db 5101 CAGGCCGACATCTGACCAATTTTGAAGCTGTCAGAGGCTGCCACTGTTTGGGACA 5160
 QY 5161 GGCAGAGTTGAGAGCGGCTCTGATCTGTGACAGCAATGCGCAGAAAGTGTCTATCTC 5220
 Db 5161 GGCAGAGTTGAGAGCGGCTCTGATCTGTGACAGCAATGCGCAGAAAGTGTCTATCTC 5220
 QY 5221 CGCTACACGAAACCTTACGAAATCTGATCCGAAAGAGATAGAGACTTCAGAGCCC 5280
 Db 5221 CGCTACACGAAACCTTACGAAATCTGATCCGAAAGAGATAGAGACTTCAGAGCCC 5280
 QY 5281 TGCAGCTGTATCCACTTCAACCAATTAAGTATCTCATTTGGAAACCAATTAATTCTAGGA 5340
 Db 5281 TGCAGCTGTATCCACTTCAACCAATTAAGTATCTCATTTGGAAACCAATTAATTCTAGGA 5340
 QY 5341 ATGACATGAGAGCTGACAGCTGAGAGAAATTCCTGATGAAGATGACCAATTCCTGGCA 5400
 Db 5341 ATGACATGAGAGCTGACAGCTGAGAGAAATTCCTGATGAAGATGACCAATTCCTGGCA 5400
 QY 5401 CTTGCTGTGTTGCGCGCTTTCACAGAGCTTCCCTGCTCATCTGACAGTGAACAGC 5460
 Db 5401 CTTGCTGTGTTGCGCGCTTTCACAGAGCTTCCCTGCTCATCTGACAGTGAACAGC 5460
 QY 5461 GCAGGCGAGAGAGAGTACTTGTGTGTTTCCAGCAATTTGGAATGTTGTTGTTCT 5520
 Db 5461 GCAGGCGAGAGAGAGTACTTGTGTGTTTCCAGCAATTTGGAATGTTGTTGTTCT 5520
 QY 5521 TACGGAAGAGCTAGCGGACAGAGATCTCAAGTGAATGCTTACCTTTGGCTTTGCC 5580
 Db 5521 TACGGAAGAGCTAGCGGACAGAGATCTCAAGTGAATGCTTACCTTTGGCTTTGCC 5580
 QY 5581 TACAGAGAACCTTATCTGTTGTGACCACTTCACTCAAGTAATTGAGATCCAG 5640

Db 5581 TACAGAGAACCTTATCTGTTGTGACCACTTCACTCACTCAAGTAATTGAGATCCAG 5640
 QY 5641 GCAGCTCTCAGAGAGGACCCCTGCGCGAGCTGACATCCGAAACCCGCGCTAC 5700
 Db 5641 GCAGCTCTCAGAGAGGACCCCTGCGCGAGCTGACATCCGAAACCCGCGCTAC 5700
 QY 5701 CTGGGCGCTGCGCTTCTTCAAGAGCAATTTACTTGGCGTCTCATACAGATTAATTA 5760
 Db 5701 CTGGGCGCTGCGCTTCTTCAAGAGCAATTTACTTGGCGTCTCATACAGATTAATTA 5760
 QY 5761 AGGTCATTTGCTGCAAGAGGAACCTGTGAAGAGTCCGCACTGAACCAACCGGGGC 5820
 Db 5761 AGGTCATTTGCTGCAAGAGGAACCTGTGAAGAGTCCGCACTGAACCAACCGGGGC 5820
 QY 5821 CCGTCCACTCCGACAGAGCCCAAGAGAGAGAGCCACCACTGAACAGACATC 5880
 Db 5821 CCGTCCACTCCGACAGAGCCCAAGAGAGAGAGCCACCACTGAACAGACATC 5880
 QY 5881 ACCAAGCGGTGCTTCAAGCCAGCGCCCGCCAGAGGCCCAACCGAGAGCCA 5940
 Db 5881 ACCAAGCGGTGCTTCAAGCCAGCGCCCGCCAGAGGCCCAACCGAGAGCCA 5940
 QY 5941 AGCACACCCCAACGCTACCGCGAGAGGCGGACCGAGCTGGCGAGGCAAGTCTTCTGCG 6000
 Db 5941 AGCACACCCCAACGCTACCGCGAGAGGCGGACCGAGCTGGCGAGGCAAGTCTTCTGCG 6000
 QY 6001 CGGCCCTGAGAGAGAGAGTCCCGCGCGAGTACTCAAGACAGCGGAGAGAGCGTCC 6060
 Db 6001 CGGCCCTGAGAGAGAGAGTCCCGCGCGAGTACTCAAGACAGCGGAGAGAGCGTCC 6060
 QY 6061 CCGCGAGGCTGTTTGAAGAAGAGAGAGGCGCGCTGCTGGGAGCCGTGAGAGAC 6120
 Db 6061 CCGCGAGGCTGTTTGAAGAAGAGAGAGGCGCGCTGCTGGGAGCCGTGAGAGAC 6120
 QY 6121 CCGGTGCTCCAGGTGAACAGGTCTGGAGACCAAGTCTTCAATATA 6165
 Db 6121 CCGGTGCTCCAGGTGAACAGGTCTGGAGACCAAGTCTTCAATATA 6165
 RESULT 4
 AAL55215
 ID AAL55215 standard; DNA; 8603 BP.
 XX
 AC AAL55215;
 XX
 DT 01-MAY-2003 (first entry)
 XX
 DE Human CRK related DNA sequence, SEQ ID No 4.
 XX
 KW Anorectic; hypotensive; cardiac; antilipemic; cerebroprotective;
 KW antiobut; osteopathic; antiarthritic; cytostatic; antidepressant;
 KW immunomodulator; antianemic; tranquiliser; antiparkinsonian; nootropic;
 KW neuroprotective; antiinflammatory; antidiabetic; analgesic;
 KW human c/lon rho/rac-interacting kinase; enzyme; CRK; ameliorating;
 KW obesity; comorbidity; cancer; anorexia; cachexia; bulimia;
 KW central nervous system disorder; chronic obstructive pulmonary disease;
 KW diabetes; pain; ds.
 XX
 OS Homo sapiens.
 PN W0200304523-A1.
 PD 16-JAN-2003.
 XX
 PF 28-JUN-2002; 2002WC-EP007156.
 XX
 PR 02-JUL-2001; 2001US-0301841P.
 PR 11-DEC-2001; 2001US-0338651P.
 PR 25-APR-2002; 2002US-0375014P.
 XX
 PA (FARB) BAYER AG.

PI Zhu Z;
XX WPI: 2003-221576/21.
DR
XX New human citron rho/rac-interacting kinase (CRIK) polypeptide and
PT polynucleotide, useful in preventing, ameliorating or treating diseases
PT associated with human CRK dysfunction, e.g. obesity, diabetes or
PT Alzheimer's disease.
XX
PS Disclosure; Fig 4; 237p; English.

XX The invention relates to an isolated polynucleotide encoding a human
CC citron rho/rac-interacting kinase polypeptide. The isolated
CC polynucleotide comprises a 6165 or 8603 base pair sequence, given in the
CC specification. The human citron rho/rac-interacting kinase (CRIK)
CC polypeptide and polynucleotide are useful in preventing, ameliorating, or
CC treating diseases associated with human CRK dysfunction such as obesity
CC and obesity-associated comorbidities (e.g. hypertension, coronary artery
CC disease, hyperlipidaemia, stroke, gout, osteoarthritis, some types of
CC cancer including endometrial, breast, prostate and colon cancer),
CC anorexia, cachexia, bulimia, central nervous system disorders (e.g. mood
CC disorders, anxiety disorders, Parkinson's disease or Alzheimer's
CC disease), chronic obstructive pulmonary disease, or diabetes. These can
CC also be used to treat pain associated with the disorders. The human CRK
CC polypeptide is also useful in diagnostic assays or in genetic testing.
CC The expression vector or the reagent is useful in preparing a medicament
CC for modulating the activity of a human CRK in a disease, e.g. obesity, a
CC central nervous system disorder, or chronic obstructive pulmonary
CC disease. The fusion protein is useful for generating antibodies against a
CC CRK polypeptide and for use in various assay systems. The methods are
CC useful in producing and detecting the polynucleotide and polypeptide and
CC in screening for agents that modulate the activity of the human CRK
CC polypeptide. This polynucleotide sequence represents a DNA sequence
CC relating to the human CRK protein of the invention
XX

Sequence 8603 BP; 2305 A; 2206 C; 2215 G; 1877 T; 0 U; 0 Other;

Query Match 99.9%; Score 6160.2; DB 9; Length 8603;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 6162; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ATGTTGAAGTTCAAAATATGAGGCGGGAATCTTTGGATGCTGTGCTGAACCAT 60
DB 1 ATGTTGAAGTTCAAAATATGAGGCGGGAATCTTTGGATGCTGTGCTGAACCAT 60
QY 61 GCCAGCCGGGCTCCAGGCTGAATCTGTTCTCCAGGGAACACCCCTTATGACTCAA 120
DB 61 GCCAGCCGGGCTCCAGGCTGAATCTGTTCTCCAGGGAACACCCCTTATGACTCAA 120
QY 121 CAGCAATGCTCTCTCTTTCCGAGAAAGGAAATATAGATCCCTCTTTGTTCTTTGAA 180
DB 121 CAGCAATGCTCTCTCTTTCCGAGAAAGGAAATATAGATCCCTCTTTGTTCTTTGAA 180
QY 181 GAATGAGTCAGCTGCTGATGAAGATTAACAGCTAGCAACTTTGTCGGAGATAT 240
DB 181 GAATGAGTCAGCTGCTGATGAAGATTAACAGCTAGCAACTTTGTCGGAGATAT 240
QY 241 TCCGACACCATAGCTAGTTACAGAGCTCCAGCTTCGGCAAAAGACTTCGAAGTCA 300
DB 241 TCCGACACCATAGCTAGTTACAGAGCTTCGAGCTTCGGCAAAAGACTTCGAAGTCA 300
QY 301 AGCTCTTGATGCTTGTGCTCACTTTGCTGAAGTCAGAGTGTAAAGAGAAACCGGG 360
DB 301 AGCTCTTGATGCTTGTGCTCACTTTGCTGAAGTCAGAGTGTAAAGAGAAACCGGG 360
QY 361 GACATATAGCTATGAAGATGAAGAGAGGCTTTATTTGGCCCAAGAGAGGTTTCA 420
DB 361 GACATATAGCTATGAAGATGAAGAGAGGCTTTATTTGGCCCAAGAGAGGTTTCA 420
QY 421 TTTTGTGAGAGAGCGGAACATATTATCTCGAAGCACAAGCCCGTGAATCCCAATTA 480
DB 421 TTTTGTGAGAGAGCGGAACATATTATCTCGAAGCACAAGCCCGTGAATCCCAATTA 480

QY 481 CAGTATGCTTTTCAGAGCAAAAATACCTTTATCTGATCATGAAATATACGCTGAGGG 540
DB 481 CAGTATGCTTTTCAGAGCAAAAATACCTTTATCTGATCATGAAATATACGCTGAGGG 540
QY 541 GACTGTGTCACCTTTGAATAGATAGAGACCAAGTTAGATGAACCTGATCACTTT 600
DB 541 GACTGTGTCACCTTTGAATAGATAGAGACCAAGTTAGATGAACCTGATCACTTT 600
QY 601 TACTAGCTGAGTATTTTGGCTGTTCACAGGCTTCATCTGATGAGTACGTGATCGA 660
DB 601 TACTAGCTGAGTATTTTGGCTGTTCACAGGCTTCATCTGATGAGTACGTGATCGA 660
QY 661 GACATCAAGCTTGAACATTTCTGTTGACCGACAGAGACATCAAGCTGTGATTTT 720
DB 661 GACATCAAGCTTGAACATTTCTGTTGACCGACAGAGACATCAAGCTGTGATTTT 720
QY 721 GSAATCGCGCGAAATGAATTCAGAACAGATGATGATGATGATGATGATGATGATG 780
DB 721 GSAATCGCGCGAAATGAATTCAGAACAGATGATGATGATGATGATGATGATGATG 780
QY 781 CCAGATTACATGCTCCTGAAAGTGTGATGATGATGATGATGATGATGATGATGATG 840
DB 781 CCAGATTACATGCTCCTGAAAGTGTGATGATGATGATGATGATGATGATGATGATG 840
QY 841 GGCCTGAGCTGTGATGCTGTGCTGATGATGATGATGATGATGATGATGATGATG 900
DB 841 GGCCTGAGCTGTGATGCTGTGCTGATGATGATGATGATGATGATGATGATGATG 900
QY 901 TCCCTGCGAGAGGAAACCTGCGCAGAACCTTCAATTAAGTTTGAATTTCCAGCG 960
DB 901 TCCCTGCGAGAGGAAACCTGCGCAGAACCTTCAATTAAGTTTGAATTTCCAGCG 960
QY 961 TTTTGAATTTTCAGATGACCCCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1020
DB 961 TTTTGAATTTTCAGATGACCCCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1020
QY 1021 TTGTTGTGGGCGCAAGAAAGAGACTGAAGTTGAAGTGTGTTGTCGCACTTTCTTC 1080
DB 1021 TTGTTGTGGGCGCAAGAAAGAGACTGAAGTTGAAGTGTGTTGTCGCACTTTCTTC 1080
QY 1081 TCTAAATTTGACTGGAACAACTGCTAATCTCTCTCTCTCTCTCTCTCTCTCTCT 1140
DB 1081 TCTAAATTTGACTGGAACAACTGCTAATCTCTCTCTCTCTCTCTCTCTCTCTCT 1140
QY 1141 TCTGAGATGACACCTCAATTTTGAATGAACAGAGAAATTCGAGGTTTCACTCTCT 1200
DB 1141 TCTGAGATGACACCTCAATTTTGAATGAACAGAGAAATTCGAGGTTTCACTCTCT 1200
QY 1201 CCGTGCAGCTGAGCCCTTCAGGCTTCTCGGGTGAAGAACTGCGGTTTGAGGTTTTCG 1260
DB 1201 CCGTGCAGCTGAGCCCTTCAGGCTTCTCGGGTGAAGAACTGCGGTTTGAGGTTTTCG 1260
QY 1261 TACAGCAAGGCACTGGGGATTTTGGTGAATCTGAGTCTGTTGTGTGGGTCTGCACTCC 1320
DB 1261 TACAGCAAGGCACTGGGGATTTTGGTGAATCTGAGTCTGTTGTGTGGGTCTGCACTCC 1320
QY 1321 CCTGCAAGACTAGCTCCATGAGAAAGAACTTCTCATCAAAAGCAAGAGCTACAGAC 1380
DB 1321 CCTGCAAGACTAGCTCCATGAGAAAGAACTTCTCATCAAAAGCAAGAGCTACAGAC 1380
QY 1381 TCTCAGACCAAGTGTCACAAGATGAGCAGAAATGACCCGGTTATCATCGAGAGTGTCA 1440
DB 1381 TCTCAGACCAAGTGTCACAAGATGAGCAGAAATGACCCGGTTATCATCGAGAGTGTCA 1440
QY 1441 GAGGTGAGGCTGTGCTTATAGTCAAGAGAGGTGAGCTGAAAGCTTCAAGTCA 1500
DB 1441 GAGGTGAGGCTGTGCTTATAGTCAAGAGAGGTGAGCTGAAAGCTTCAAGTCA 1500
QY 1501 TCCCTTCGAGAGCAGACCTTGTCTACTTACATCAAGATGAGTGAAGCTTAAAGAGAGT 1560
DB 1501 TCCCTTCGAGAGCAGACCTTGTCTACTTACATCAAGATGAGTGAAGCTTAAAGAGAGT 1560
QY 1561 TTGAGCAAGACCGAGTGAAGGTGTCCAGAGAGATGACAAAGCACTGCAAGCTTCTCAT 1620

```
Db 1561 TTGAGAGACAGATGAGGTGTCCAGAGAGATGACAAAGACATGCGCTTCTCCAT 1620
Qy 1621 GATATCAGAGAGAGAGCCGGAAGCTCCAGAAATCAAAGAGAGATGACAGGCTCAA 1680
Db 1621 GATATCAGAGAGAGAGCCGGAAGCTCCAGAAATCAAAGAGAGATGACAGGCTCAA 1680
Qy 1681 GTGGAGAAATGAGGTTGATGATGAATCAAGTTGAAAGAGATCTTGTCTCAGCAAGAA 1740
Db 1681 GTGGAGAAATGAGGTTGATGATGAATCAAGTTGAAAGAGATCTTGTCTCAGCAAGAA 1740
Qy 1741 CGAGAGATCTTCAACGAATCTGAGAGATCTTCGGCTTCTGTGTAAGAAATTCAG 1800
Db 1741 CGAGAGATCTTCAACGAATCTGAGAGATCTTCGGCTTCTGTGTAAGAAATTCAG 1800
Qy 1801 CGGAAAGCGACAGAAATGTGACATTAATCTGTTGAAGCTTAAGATCAAGGAAAGCTGAA 1860
Db 1801 CGGAAAGCGACAGAAATGTGACATTAATCTGTTGAAGCTTAAGATCAAGGAAAGCTGAA 1860
Qy 1861 GTGGAGAAATATCGAAATCTGAGAAAGATCAATGCTGAGCAGCAGCTCAAAATTCAGAG 1920
Db 1861 GTGGAGAAATATCGAAATCTGAGAAAGATCAATGCTGAGCAGCAGCTCAAAATTCAGAG 1920
Qy 1921 CTCGAGAGAAATCTGAGAGAGGCTGTAAAAGCCAGCAGGAGCCAGGCTGCTGAG 1980
Db 1921 CTCGAGAGAAATCTGAGAGAGGCTGTAAAAGCCAGCAGGAGCCAGGCTGCTGAG 1980
Qy 1981 AATATCCGCGAGCAAGAGAGAGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2040
Db 1981 AATATCCGCGAGCAAGAGAGAGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2040
Qy 2041 GATTTCTTGAAGGATCAGAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2100
Db 2041 GATTTCTTGAAGGATCAGAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2100
Qy 2101 GAGAAACAGGTAAGAGAGATGAGAGCATTGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2160
Db 2101 GAGAAACAGGTAAGAGAGATGAGAGCATTGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2160
Qy 2161 ATCCAGACAAATATCCAGACAGATCCAGCAGATGAGCTGATTAATTTCTGAGGCTGAGAG 2220
Db 2161 ATCCAGACAAATATCCAGACAGATCCAGCAGATGAGCTGATTAATTTCTGAGGCTGAGAG 2220
Qy 2221 AAACATCGGAGAGCCCAAGTCTCAGCCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 2280
Db 2221 AAACATCGGAGAGCCCAAGTCTCAGCCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 2280
Qy 2281 CAGCACTATGAGAGAAAGATTAAAGTTGAGCAATCAATCAATCAATCAATCAATCAATCA 2340
Db 2281 CAGCACTATGAGAGAAAGATTAAAGTTGAGCAATCAATCAATCAATCAATCAATCAATCA 2340
Qy 2341 AAGGAGACATGAGAGAAATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2400
Db 2341 AAGGAGACATGAGAGAAATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2400
Qy 2401 ATTCTCAGGAGAGAGAGAGAGAGATGATCAATGATGATGATGATGATGATGATGATGAT 2460
Db 2401 ATTCTCAGGAGAGAGAGAGAGAGATGATCAATGATGATGATGATGATGATGATGATGAT 2460
Qy 2461 CAGAGAGATTTGGAAGTGTCTGAGAGCAGATTAATCTTCAGCAGAAATGAGCAGCTTTTAC 2520
Db 2461 CAGAGAGATTTGGAAGTGTCTGAGAGCAGATTAATCTTCAGCAGAAATGAGCAGCTTTTAC 2520
Qy 2521 CAAAGAGACATGAGAGAGAGAGAGAGATGATTTGAGACTTCAGAGCAGAGAGAGAGAG 2580
Db 2521 CAAAGAGACATGAGAGAGAGAGAGAGATGATTTGAGACTTCAGAGCAGAGAGAGAGAG 2580
Qy 2581 CTGAGAGACAGAGCTGGAGAGTTGAGAGCCAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2640
Db 2581 CTGAGAGACAGAGCTGGAGAGTTGAGAGCCAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2640
Qy 2641 AAGATCAGCCACCAAGACCAAGTGAACAAGATCGGCTGCTGGAATCTGAGAGAGAGTTG 2700
Db 2641 AAGATCAGCCACCAAGACCAAGTGAACAAGATCGGCTGCTGGAATCTGAGAGAGAGTTG 2700
Db 2641 AAGATCAGCCACCAAGACCAAGTGAACAAGATCGGCTGCTGGAATCTGAGAGAGAGTTG 2700
Qy 2701 CGGAGAGTCAAGTTAGAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2760
Db 2701 CGGAGAGTCAAGTTAGAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2760
Qy 2761 CTACAGCTCTCCCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2820
Db 2761 CTACAGCTCTCCCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2820
Qy 2821 GCCCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2880
Db 2821 GCCCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2880
Qy 2881 GAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2940
Db 2881 GAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2940
Qy 2941 CTTCTGTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3000
Db 2941 CTTCTGTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3000
Qy 3001 GACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3060
Db 3001 GACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3060
Qy 3061 GCGCGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3120
Db 3061 GCGCGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3120
Qy 3121 ACGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3180
Db 3121 ACGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3180
Qy 3181 TGACCAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3240
Db 3181 TGACCAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3240
Qy 3241 AAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3300
Db 3241 AAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3300
Qy 3301 TGTGAGGTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3360
Db 3301 TGTGAGGTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3360
Qy 3361 GATCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3420
Db 3361 GATCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3420
Qy 3421 GAGATTCGAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3480
Db 3421 GAGATTCGAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3480
Qy 3481 GACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3540
Db 3481 GACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3540
Qy 3541 CAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3600
Db 3541 CAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3600
Qy 3601 TTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3660
Db 3601 TTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3660
Qy 3661 GAGAGCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3720
Db 3661 GAGAGCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3720
Qy 3721 GAAAGATTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3780
Db 3721 GAAAGATTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3780
```



```
QY 3781 CAAACCAACTCATTTGATTTTCTGCAAGCCAAATGGAACAACCTGCTTAAAAAAGAAAAG 3840
Db 3781 CAAACCAACTCATTTGATTTTCTGCAAGCCAAATGGAACAACCTGCTTAAAAAAGAAAAG 3840
QY 3841 GTTCTCTGAGTACATAGAGTGAAGCTGGCCCTGGAGAAAGGAAAGCTGGCTGCA 3900
Db 3841 GTTCTCTGAGTACATAGAGTGAAGCTGGCCCTGGAGAAAGGAAAGCTGGCTGCA 3900
QY 3901 GAGCTAGAGAGACCCCTTCAAGAAACCCGATCGAGCTCCGGTCCGGCCGGAGAGAGCT 3960
Db 3901 GAGCTAGAGAGACCCCTTCAAGAAACCCGATCGAGCTCCGGTCCGGCCGGAGAGAGCT 3960
QY 3961 GCCCACCAGAAAGCAAGCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 4020
Db 3961 GCCCACCAGAAAGCAAGCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 4020
QY 4021 ATGCGCATGTCGCGCATTCGTCGCGGTCGCGGAGACACAGCCGATGCGATGAGCTGCTG 4080
Db 4021 ATGCGCATGTCGCGCATTCGTCGCGGTCGCGGAGACACAGCCGATGCGATGAGCTGCTG 4080
QY 4081 GCCCGGCATTCACGCGCAGAAAGAGTCTTCAACTCCAGAGAAATTTAGTCGGCGCTT 4140
Db 4081 GCCCGGCATTCACGCGCAGAAAGAGTCTTCAACTCCAGAGAAATTTAGTCGGCGCTT 4140
QY 4141 AAGGAAACGATGACACATATTTCTCACCGATTCAAGCTAGAGCTGAACATGCAAGCC 4200
Db 4141 AAGGAAACGATGACACATATTTCTCACCGATTCAAGCTAGAGCTGAACATGCAAGCC 4200
QY 4201 ACAAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4260
Db 4201 ACAAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4260
QY 4261 GAATGTCAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4320
Db 4261 GAATGTCAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4320
QY 4321 CCTGCTGAATATGTGCAACATTTACCGAGGCTTTCTGCGGTGACAAATGAACTCCCA 4380
Db 4321 CCTGCTGAATATGTGCAACATTTACCGAGGCTTTCTGCGGTGACAAATGAACTCCCA 4380
QY 4381 GGTCTCCAGACCAAGAGCCGACGACGCTTGTGCACTTGTGAAGGAGGTGTGAAGTGTG 4440
Db 4381 GGTCTCCAGACCAAGAGCCGACGACGCTTGTGCACTTGTGAAGGAGGTGTGAAGTGTG 4440
QY 4441 AAGAAATACAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4500
Db 4441 AAGAAATACAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4500
QY 4501 AAAAGTCTCATTTATGACATGAGCAGAGAGCTGGAAGAGCTGGAAGAGAGAGAGAG 4560
Db 4501 AAAAGTCTCATTTATGACATGAGCAGAGAGCTGGAAGAGCTGGAAGAGAGAGAGAG 4560
QY 4561 GAGCTGTGCTTCCGAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4620
Db 4561 GAGCTGTGCTTCCGAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4620
QY 4621 GCAAAATACAGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4680
Db 4621 GCAAAATACAGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4680
QY 4681 ACCTGCTGAGCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4740
Db 4681 ACCTGCTGAGCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4740
QY 4741 TGGGTCAACGCGCTTGAAGATCAATTTGTCCGAGGTGGAGAGGTTTCTAGAGAGAGAG 4800
Db 4741 TGGGTCAACGCGCTTGAAGATCAATTTGTCCGAGGTGGAGAGGTTTCTAGAGAGAGAG 4800
QY 4801 GCTGATGCTAAACTGCTTGGAAATCTCCTGTGAAATCTGAAAGGTGAATGACCTTGAAC 4860
Db 4801 GCTGATGCTAAACTGCTTGGAAATCTCCTGTGAAATCTGAAAGGTGAATGACCTTGAAC 4860
```

```
QY 4861 ATGACTGCAACGCTGCCCTTCACTGACAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4920
Db 4861 ATGACTGCAACGCTGTGCTTCACTGACAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4920
QY 4921 TAGCCCTGTAATGTCTTGAAGAAATCTCCCTAACCATGTCCAGAGATTTGAGAGATCTT 4980
Db 4921 TAGCCCTGTAATGTCTTGAAGAAATCTCCCTAACCATGTCCAGAGATTTGAGAGATCTT 4980
QY 4981 CAAATTTATATATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5040
Db 4981 CAAATTTATATATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5040
QY 5041 CTGTGTCTTGTGAGAGTGAAGAAAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5100
Db 5041 CTGTGTCTTGTGAGAGTGAAGAAAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5100
QY 5101 CAGCCGCAATCTCAACCAACATTTTGAAGAGTGAAGAGAGAGAGAGAGAGAGAGAGAG 5160
Db 5101 CAGCCGCAATCTCAACCAACATTTTGAAGAGTGAAGAGAGAGAGAGAGAGAGAGAGAG 5160
QY 5161 GGCAGAGATTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5220
Db 5161 GGCAGAGATTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5220
QY 5221 CGCTACAGAGAAACCTCAGCAAAATATGCAATCCGAGAGAGAGAGAGAGAGAGAGAG 5280
Db 5221 CGCTACAGAGAAACCTCAGCAAAATATGCAATCCGAGAGAGAGAGAGAGAGAGAGAG 5280
QY 5281 TGCAAGCTGATCACTTCAACCAATTAAGATCTGATCTGATGAAGAGAGAGAGAGAGAG 5340
Db 5281 TGCAAGCTGATCACTTCAACCAATTAAGATCTGATCTGATGAAGAGAGAGAGAGAGAG 5340
QY 5341 ATGACATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5400
Db 5341 ATGACATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5400
QY 5401 CCTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 5460
Db 5401 CCTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 5460
QY 5461 GCAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5520
Db 5461 GCAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5520
QY 5521 TACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5580
Db 5521 TACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5580
QY 5581 TACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5640
Db 5581 TACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5640
QY 5641 GAGAGCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5700
Db 5641 GAGAGCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5700
QY 5701 CTGAGGCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5760
Db 5701 CTGAGGCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5760
QY 5761 AAGGTCAATTTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5820
Db 5761 AAGGTCAATTTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5820
QY 5821 CGGTCAACCTCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5880
Db 5821 CGGTCAACCTCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5880
QY 5881 ACCAAGCGGTGAGCTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5940
Db 5881 ACCAAGCGGTGAGCTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5940
QY 5941 AGCAACCCCAACCGCTACCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6000
```

Db 5941 AGACACCCACCGCTACCGCGAGGGCGGACCGAGCTGGCAGGACAGTCTCTCGGC 6000
QY CGCCCTGGAGGAGAGAGTCCCGCGCGGATCTAGCAGCGGAGAGCCGCTCC 6060
Db 6001 CGCCCTGGAGGAGAGTCCCGCGCGGATCTAGCAGCGGAGAGCCGCTCC 6060
QY 6061 CGCCCTGGAGGAGAGTCCCGCGCGGATCTAGCAGCGGAGAGCCGCTCC 6060
Db 6061 CGCCCTGGAGGAGAGTCCCGCGCGGATCTAGCAGCGGAGAGCCGCTCC 6060
QY 6061 CGCCCTGGAGGAGAGTCCCGCGCGGATCTAGCAGCGGAGAGCCGCTCC 6060
Db 6061 CGCCCTGGAGGAGAGTCCCGCGCGGATCTAGCAGCGGAGAGCCGCTCC 6060
QY 6121 CGCCTGTCCGAGTGAACAGAGTCTGGGACCACTTCTCACTATA 6165
Db 6121 CGCCTGTCCGAGTGAACAGAGTCTGGGACCACTTCTCACTATA 6165
QY 6121 CGCCTGTCCGAGTGAACAGAGTCTGGGACCACTTCTCACTATA 6165
Db 6121 CGCCTGTCCGAGTGAACAGAGTCTGGGACCACTTCTCACTATA 6165
RESULT 5
ADJ96544
ID ADJ96544 standard; DNA; 8656 BP.
XX ADJ96544;
AC
XX 06-MAY-2004 (first entry)
DT
XX
DE Human citron Rho-interacting kinase CRK DNA Seg1D1.
XX
XX Gene; de; kinase; human; SNP; single nucleotide polymorphism;
KM tyrosine protein kinase; serine/threonine protein kinase; PTX; STK;
KM gene therapy; cancer; immune-related disease; cardiovascular disease;
KM brain; neuronal associated disease; metabolic; inflammatory disorder;
KM cytosolic; neuroprotective; immunomodulator; antiinflammatory;
KM citron Rho-interacting kinase; CRK.
XX
OS Homo sapiens.
OS 51.
XX
XX Key Location/Qualifiers
FH replace(2924,g)
FT /*tag= a
FT /standard_name= "Single nucleotide polymorphism"
FT variation replace(2933,c)
FT /*tag= b
FT /standard_name= "Single nucleotide polymorphism"
FT variation replace(3377,a)
FT /*tag= c
FT /standard_name= "Single nucleotide polymorphism"
FT variation replace(4085,c)
FT /*tag= d
FT /standard_name= "Single nucleotide polymorphism"
FT variation replace(7676,c)
FT /*tag= e
FT /standard_name= "Single nucleotide polymorphism"
XX
XX MO2004006838-A2.
XX
XX 22-JAN-2004.
XX
XX 15-JUL-2003; 2003MO-US021730.
XX
XX 15-JUL-2002; 2002US-0395632P.
XX
XX (SUGEN-) SUGEN INC.
XX
XX Whyte D, Manning G, Caenepeel S;
XX
XX MPI; 2004-122753/12.
XX P-PSDB; ADJ96610.
XX
XX New nucleic acid molecule encoding a kinase polypeptide, useful for
PT preparing a composition for treating diseases or disorders, e.g., cancer,
PT or neurological, immunological or inflammatory disorders.
XX
XX Example 1; SEQ ID NO 1; 366bp; English.

XX This invention relates to a novel isolated, enriched or purified nucleic
CC acid molecule that encodes a kinase polypeptide. Specifically, it relates
CC to human tyrosine and serine/threonine protein kinases (PTK's and STK's),
CC as well as protein kinase-like enzymes. The present invention describes
CC screening methods to identify agonists, antagonists and antibodies that
CC can be used to modulate the activity or function of the mammalian kinase
CC enzymes. As such, these compositions can be used for gene therapy
CC purposes to treat diseases or disorders including cancer, immune-related
CC diseases, cardiovascular disease, brain or neuronal associated disease,
CC metabolic and inflammatory disorders. Accordingly, they exhibit
CC cytosolic, neuroprotective, immunomodulator and antiinflammatory
CC activities. This polynucleotide sequence is a human kinase DNA sequence
CC of the invention.
XX
SQ Sequence 8656 BP; 2314 A; 2219 C; 2239 G; 1884 T; 0 U; 0 Other;
Query Match 99.7%; Score 6144; DB 12; Length 8656;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 6160; Conservative 0; Mismatches 5; Indels 3; Gaps 1;
QY 1 ATGTGAAGTCAAAATATGAGGCGGAAATCCTTTGGATGCTGTGCTGAACCAT 60
Db 51 ATGTGAAGTCAAAATATGAGGCGGAAATCCTTTGGATGCTGTGCTGAACCAT 110
QY 61 GCCAGCCGGGCTCCAGGCTGAATCTGTTCTTCAGGGGAAACACCTTTATGACTCAA 120
Db 111 GCCAGCCGGGCTCCAGGCTGAATCTGTTCTTCAGGGGAAACACCTTTATGACTCAA 170
QY 121 CAGCAGATGCTCTCTCTTCCGAGAGAGGATTTAGATGCTCTTGTCTTTGAA 180
Db 171 CAGCAGATGCTCTCTCTTCCGAGAGAGGATTTAGATGCTCTTGTCTTTGAA 230
QY 181 GAATGACATGACGCTCTGATGAAGATTAACAGTGAACCTTGTCCGGA---G 237
Db 231 GAATGACATGACGCTCTGATGAAGATTAACAGTGAACCTTGTGTCCGGAAGTG 290
QY 238 TATTCGACACATAGCTAGTTACAGAGCTCCAGCTTCGGCAAGACTTCGAATGC 297
Db 291 TATTCGACACATAGCTAGTTACAGAGCTCCAGCTTCGGCAAGACTTCGAATGC 350
QY 298 AGAAGTCTGTAGTGTGTGCTCACTTGTCTGAAGTGCAGGTGTGAAGAAAGCAAC 357
Db 351 AGAAGTCTGTAGTGTGTGCTCACTTGTCTGAAGTGCAGGTGTGAAGAAAGCAAC 410
QY 358 GGGGACATCTAGTCTATGAAGATGAAGAAAGGCTTTATTTGGCCAGAGCAGGT 417
Db 411 GGGGACATCTAGTCTATGAAGATGAAGAAAGGCTTTATTTGGCCAGAGCAGGT 470
QY 418 TCAATTTTGAAGAGCGGAACATATTATCTCGAAGCAAGCCGTGATCCCCAA 477
Db 471 TCAATTTTGAAGAGCGGAACATATTATCTCGAAGCAAGCCGTGATCCCCAA 530
QY 478 TTACAGTATGCTTTGAGGCAAAATACCTTATCTGTGCTATGAAATACGCTTGA 537
Db 531 TTACAGTATGCTTTGAGGCAAAATACCTTATCTGTGCTATGAAATACGCTTGA 590
QY 538 GGGGACTTGTGCTCACTTTGAAATAGATAGAGCAAGTTAATGAAACCTGATCAG 597
Db 591 GGGGACTTGTGCTCACTTTGAAATAGATAGAGCAAGTTAATGAAACCTGATCAG 650
QY 598 TTTTACCTAGCTGAGTGAATTTTGGCTGTCAAGCGTTTATCTGATGAGTACGTGAT 657
Db 651 TTTTACCTAGCTGAGTGAATTTTGGCTGTCAAGCGTTTATCTGATGAGTACGTGAT 710
QY 658 CGAGACATCAAGCTGAGAAATCTCTGTTGACCGGACAGAGACATCAAGCTGTGAT 717
Db 711 CGAGACATCAAGCTGAGAAATCTCTGTTGACCGGACAGAGACATCAAGCTGTGAT 770
QY 718 TTTGATCTGCGCGGAAATGAATTTCAACAAATGATGATCCAACTCCGATTTGG 777
Db 771 TTTGATCTGCGCGGAAATGAATTTCAACAAATGATGATCCAACTCCGATTTGG 830

QY 778 ACCCCAGATTACATGCGTCCGTGAAGTGCTGACTGTGATGAACGGGGATGGAAGGACACC 837
DB 831 ACCCCAGATTACATGCGTCCGTGAAGTGCTGACTGTGATGAACGGGGATGGAAGGACACC 890
QY 838 TAAGCGCTGGAAGCTGTGACTGTGAGTGTGAGTGGCGTGTGATGAGTGAATGATTTATGGG 897
DB 891 TAAGCGCTGGAAGCTGTGACTGTGAGTGTGAGTGGCGTGTGATGAGTGAATGATTTATGGG 950
QY 898 AGATGCCCTTCGCAAGGGAACCTGTGCGAAGCCTTCAATTAACATTTATGATTTTCAG 957
DB 951 AGATGCCCTTCGCAAGGGAACCTGTGCGAAGCCTTCAATTAACATTTATGATTTTCAG 1010
QY 958 CGGTTTTGAAATTTCCAGATGACCCCAAGTGAAGGATGACTTTCTGTGATTCGATTCAA 1017
DB 1011 CGGTTTTGAAATTTCCAGATGACCCCAAGTGAAGGATGACTTTCTGTGATTCGATTCAA 1070
QY 1018 AGCTGTGTGCGGCCGAAGAGAGACTGAAGTTTGAAGTCTTTGCTGCCATCTTTC 1077
DB 1071 AGCTGTGTGCGGCCGAAGAGAGACTGAAGTTTGAAGTCTTTGCTGCCATCTTTC 1130
QY 1078 TTTCTTAAATTTGACTGGAACAACTTCGTAACTCTCTCCCTCCCTTGTTCCACCTTC 1137
DB 1131 TTTCTTAAATTTGACTGGAACAACTTCGTAACTCTCTCCCTCCCTTGTTCCACCTTC 1190
QY 1138 AAGCTGAGATGACACCTCCAAATTTGATGAACCGAGAGAAATTCGTGGTTTCAATCC 1197
DB 1191 AAGCTGAGATGACACCTCCAAATTTGATGAACCGAGAGAAATTCGTGGTTTCAATCC 1250
QY 1198 TCTCGGTGCGAGTGAAGCCCTCAGGCTTCTCGGTGGAAGACTGCCGTTTGTGGGGTTT 1257
DB 1251 TCTCGGTGCGAGTGAAGCCCTCAGGCTTCTCGGTGGAAGACTGCCGTTTGTGGGGTTT 1310
QY 1258 TCGTAAAGCAAGGCACTGGGGATTTCTTGGTAACTGAGTCTGTGTGTGCTGGAC 1317
DB 1311 TCGTAAAGCAAGGCACTGGGGATTTCTTGGTAACTGAGTCTGTGTGTGCTGGAC 1370
QY 1318 TCCCTGCGCAAGCTAGCTCCATGGAAGAACTTCTCATCAAAAGCAAGGCTTACAA 1377
DB 1371 TCCCTGCGCAAGCTAGCTCCATGGAAGAACTTCTCATCAAAAGCAAGGCTTACAA 1430
QY 1378 GACTCTCAGGACAAAGTGTCAACAATGAGCAGAGAAATGACCCGGTTAATCTGAGAGTG 1437
DB 1431 GACTCTCAGGACAAAGTGTCAACAATGAGCAGAGAAATGACCCGGTTAATCTGAGAGTG 1490
QY 1438 TCAAGGTGAGGCTGTGCTTAAGTCAAGAGAGGTGAGTGAAGGCTTCTGAGACTCAG 1497
DB 1491 TCAAGGTGAGGCTGTGCTTAAGTCAAGAGAGGTGAGTGAAGGCTTCTGAGACTCAG 1550
QY 1498 AGATCCCTCTGAGAGAGACCTTGTACTTACATCAAGAAATGCAATGCTTAAAGCA 1557
DB 1551 AGATCCCTCTGAGAGAGACCTTGTACTTACATCAAGAAATGCAATGCTTAAAGCA 1610
QY 1558 AGTTTGAACAAGCAGAGTGAAGTGTCCAGAGAGATGACAAAGCACTGACGCTTCTC 1617
DB 1611 AGTTTGAACAAGCAGAGTGAAGTGTCCAGAGAGATGACAAAGCACTGACGCTTCTC 1670
QY 1618 CATGATATCAGAGAGAGAGCCGGAAAGCTTCAAAGAAATCAAGAGCAGAGATGCCAGCT 1677
DB 1671 CATGATATCAGAGAGAGAGCCGGAAAGCTTCAAAGAAATCAAGAGCAGAGATGCCAGCT 1730
QY 1678 CAAGTGAAGAAATGAGTGAATGAATCAAGTGAAGAGATTTGTCTCAAGAAATA 1737
DB 1731 CAAGTGAAGAAATGAGTGAATGAATCAAGTGAAGAGATTTGTCTCAAGAAATA 1790
QY 1738 AGACGGAATGATCTCTACGAATCTGAGCTGAGAGAGTCTGGCTTGTCTGTGAAGATTC 1797
DB 1791 AGACGGAATGATCTCTACGAATCTGAGCTGAGAGAGTCTGGCTTGTCTGTGAAGATTC 1850
QY 1798 AAGCGGAAAGCAGAGATGTCAAGATTAACCTGTTGAAGGCTTAAGATCAAGAGAGCCT 1857
DB 1851 AAGCGGAAAGCAGAGATGTCAAGATTAACCTGTTGAAGGCTTAAGATCAAGAGAGCCT 1910
QY 1858 GAAAGTGGAGAAATATCCGAAACTGAGAGAAATCAATGCTGAGCAGACCTCAAAATTCAG 1917

DB 1911 GAAAGTGGAGAAATATGAGAACTGAGAGAAATCAATGCTGAGCAGACCTCAAAATTCAG 1970
QY 1918 GAGCTCCAGAGAACTGAGAAAGGCTGTAAAGCCAGACAGAGGCCACCGAGCTGCTG 1977
DB 1971 GAGCTCCAGAGAACTGAGAAAGGCTGTAAAGCCAGACAGAGGCCACCGAGCTGCTG 2030
QY 1978 CAGAAATTCGCGAGGCAAGAGCGAGCCGAGAGGGAGCTGGAAGATGTGCAAGCCGA 2037
DB 2031 CAGAAATTCGCGAGGCAAGAGCGAGCCGAGAGGGAGCTGGAAGATGTGCAAGCCGA 2090
QY 2038 GAGATTCCTTCTGAAGGATCAAGAAAGAGCTGTGAAGCTGAAGAACGCCCATCTTCT 2097
DB 2091 GAGATTCCTTCTGAAGGATCAAGAAAGAGCTGTGAAGCTGAAGAACGCCCATCTTCT 2150
QY 2098 CTGAGAACCAAGGTAAAGAGACTAGAGACCATGAGCGTGAAGAAAACAGACTGAAGAT 2157
DB 2151 CTGAGAACCAAGGTAAAGAGACTAGAGACCATGAGCGTGAAGAAAACAGACTGAAGAT 2210
QY 2158 GACATCCAGCAAAATCCCAAGATCCAGAGATGAGTGAATTAATTTTGAAGCTCGAA 2217
DB 2211 GACATCCAGCAAAATCCCAAGATCCAGAGATGAGTGAATTAATTTTGAAGCTCGAA 2270
QY 2218 GAGAAACATCGGAGGCCCAAGTCTGAGCCAGCACTTGAAGTGAACCTGAAACAGAAA 2277
DB 2271 GAGAAACATCGGAGGCCCAAGTCTGAGCCAGCACTTGAAGTGAACCTGAAACAGAAA 2330
QY 2278 GAGCAGACTATGAGAAAGATTAAGTGTGGAACAATCAGATTAAGAAAGACTGGCT 2337
DB 2331 GAGCAGACTATGAGAAAGATTAAGTGTGGAACAATCAGATTAAGAAAGACTGGCT 2390
QY 2338 GACAGAGAGCACTGAGAAACATGATCAGAGACACAGAGAGAGGCCCATGAAAGGCG 2397
DB 2391 GACAGAGAGCACTGAGAGAAACATGATCAGAGACACAGAGAGAGGCCCATGAAAGGCG 2450
QY 2398 AAAATTCAGGAGAAACAGAGGGATGATCAATGCTATGGAATTCGAAGATTCAGATCCCTG 2457
DB 2451 AAAATTCAGGAGAAACAGAGGGATGATCAATGCTATGGAATTCGAAGATTCAGATCCCTG 2510
QY 2458 GAAACAGAGATTTGTGAACTGTCTGAAAGCCAAATTAACCTGACAGCAATATGACGCTTTT 2517
DB 2511 GAAACAGAGATTTGTGAACTGTCTGAAAGCCAAATTAACCTGACAGCAATATGACGCTTTT 2570
QY 2518 ACCCAAGGAAATGAGGCCCAAGAGAGATGATTTCTGAATCTCAGGCAACAGAAATTT 2577
DB 2571 ACCCAAGGAAATGAGGCCCAAGAGAGATGATTTCTGAATCTCAGGCAACAGAAATTT 2630
QY 2578 TAACCTGAGACACAGGCTGGGAAGTTGAGAGGCCCAAGAACCGGAAATCTGAGAGAGACTG 2637
DB 2631 TAACCTGAGACACAGGCTGGGAAGTTGAGAGGCCCAAGAACCGGAAATCTGAGAGAGACTG 2690
QY 2638 GAGAAAGATCAGCCCAAGACCAAGTGAACAAGATTCGGCTCTGGAATCTGAGACAAAGA 2697
DB 2691 GAGAAAGATCAGCCCAAGACCAAGTGAACAAGATTCGGCTCTGGAATCTGAGACAAAGA 2750
QY 2698 TTTGCGGAGGTCAAGTCTAGAGACAGAGAGCAAGAACTGAGACTCAAGCCGCACTCACA 2757
DB 2751 TTTGCGGAGGTCAAGTCTAGAGACAGAGAGCAAGAACTGAGACTCAAGCCGCACTCACA 2810
QY 2758 GAGCTACAGCTCTCCCTGAGAGAGGCGAGTCAAGTTGAACAGCCCTGAGAGCTCAGCG 2817
DB 2811 GAGCTACAGCTCTCCCTGAGAGAGGCGAGTCAAGTTGAACAGCCCTGAGAGCTCAGCG 2870
QY 2818 GCGGCCCTGAGAGAGCACTTGGCAGAGCGAAGACAGAGCTGGAAGAGACCAACAGAGA 2877
DB 2871 GCGGCCCTGAGAGAGCACTTGGCAGAGCGAAGAGAGCTGGAAGAGAGACCAACAGAGA 2930
QY 2878 GCTGAAGAGAGATCAGAGCACTCAGGCACTAAGAGATGAATTCAGCGCAAAATTTGAT 2937
DB 2931 GCTGAAGAGAGATCAGAGCACTCAGGCACTAAGAGATGAATTCAGCGCAAAATTTGAT 2990
QY 2938 GCTCTTGTAAACAGCTGTAACTGTAACTCAAGACCTGAGAGAGAGACTTAAACAGCTGACC 2997

Db	2991	GCTCTTGTAACAGCTGTACTGTATATCAACAAGCTGGAGAGACGTAAACCACTGACC	3050
Oy	2998	GAGGACAAACGGTGAATCAACAAACCAAACTTCTACTGTGCCAAACAATCGATGAGGCT	3057
Db	3051	GAGGACAAACGGTGAATCAACAAACCAAACTTCTACTGTGCCAAACAATCGATGAGGCT	3110
Oy	3058	TCTGGCCGCAACGACGAGATTGTACACTGCGGAAGTGAATGACATCTCCCGGGAG	3117
Db	3111	TCTGGCCGCAACGACGAGATTGTACACTGCGGAAGTGAATGACATCTCCCGGGAG	3170
Oy	3118	ATCACGGAACGAGAGATGACAGCTTACAGCCAGAAACAAAGATGGAGGCTCTGAAGAC	3177
Db	3171	ATCACGGAACGAGAGATGACAGCTTACAGCCAGAAACAAAGATGGAGGCTCTGAAGAC	3230
Oy	3178	ACGTGCAACATGCTGGAGAAACAGTTCATGGAATTTGGAGGCCCTAAACGATGAGCTGTA	3237
Db	3231	ACGTGCAACATGCTGGAGAAACAGTTCATGGAATTTGGAGGCCCTAAACGATGAGCTGTA	3290
Oy	3238	GAAAAAGAGCGGCACTGGGAGGCGCTTGGAGAGCGTCTTGSGTATGAGAAATCCCATTT	3297
Db	3291	GAAAAAGAGCGGCACTGGGAGGCGCTTGGAGAGCGTCTTGSGTATGAGAAATCCCATTT	3350
Oy	3298	GAGTGTCCGGTTCGAGAGCTGCAGAGATGCTGCAACCCGAGAAACAGAGAGGGCGAGA	3357
Db	3351	GAGTGTCCGGTTCGAGAGCTGCAGAGATGCTGCAACCCGAGAAACAGAGAGGGCGAGA	3410
Oy	3358	GCCGATCAGCGGATATCACCGAGTCTTCGACAGGTGGTGGAGCTGGGCGATGAAGACACAG	3417
Db	3411	GCCGATCAGCGGATATCACCGAGTCTTCGACAGGTGGTGGAGCTGGGCGATGAAGACACAG	3470
Oy	3418	GCTGAGATTCTGCTCTGCACAGGCTCTCAABAGACAGAGCTGAAGGCCGAGAGCCTC	3477
Db	3471	GCTGAGATTCTGCTCTGCACAGGCTCTCAABAGACAGAGCTGAAGGCCGAGAGCCTC	3530
Oy	3478	TCTGCAAGCTCATTAGACCTGAGAAAGAGATGCTAGCTTGAATGAATGAATGCCGAAGC	3537
Db	3531	TCTGCAAGCTCATTAGACCTGAGAAAGAGATGCTAGCTTGAATGAATGAATGCCGAAGC	3590
Oy	3538	TTACAGCAGAAAGCTGGAGACTGAACGAGAGCTCAACAGAGGCTTCTGGAAGACCAAGC	3597
Db	3591	TTACAGCAGAAAGCTGGAGACTGAACGAGAGCTCAACAGAGGCTTCTGGAAGACCAAGC	3650
Oy	3658	CAGAAGCTCTAGATCGGGCTGATCTACTGAAGACAGAAAGATGACTTGGAGTATCAG	3717
Db	3711	CAGAAGCTCTAGATCGGGCTGATCTACTGAAGACAGAAAGATGACTTGGAGTATCAG	3770
Oy	3718	CTGGAAAAATTCCAGGTTCTCTAATTCTCATGAAAGGTGAATGGAGGACATATTTCT	3777
Db	3771	CTGGAAAAATTCCAGGTTCTCTAATTCTCATGAAAGGTGAATGGAGGACATATTTCT	3830
Oy	3778	CAACAAACCAACTCATTGATTTTCTGCAGCGCAAAATGAGACCAACTGCTAAAAAGAA	3837
Db	3831	CAACAAACCAACTCATTGATTTTCTGCAGCGCAAAATGAGACCAACTGCTAAAAAGAA	3890
Oy	3838	AAAGTCTCTCTGCACTAACAATGAGCTGAAGCTGCGCTTGGAGAAAGAGAAAGCTCGCTGT	3897
Db	3891	AAAGTCTCTCTGCACTAACAATGAGCTGAAGCTGCGCTTGGAGAAAGAGAAAGCTCGCTGT	3950
Oy	3898	GCAGAGCTAGAGGAAGCCCTTCAGAAAGACCCGATCGAGCTCCGGTCCGCCCGGAGAGGA	3957
Db	3951	GCAGAGCTAGAGGAAGCCCTTCAGAAAGACCCGATCGAGCTCCGGTCCGCCCGGAGAGGA	4010
Oy	3958	GCTGCGCAACCGCAAGCAACGAGACCAACCAACCCATCAGCGCCAGCCAGCGAGCAG	4017
Db	4011	GCTGCGCAACCGCAAGCAACGAGACCAACCAACCCATCAGCGCCAGCCAGCGAGCAG	4070
Oy	4018	CAGATCGCATATGTCGGCATATGTCGGTGTGCGCAAGACACAGCCCATGTCATGAGCTTG	4077
Db	4071	CAGATCGCATATGTCGGCATATGTCGGTGTGCGCAAGACACAGCCCATGTCATGAGCTTG	4130

QY	4078	CTGACCCTGGCATCCAGCCGAGAAAGAGTCTTCACTCCAGAGAAATTTAGTCGGCGT	4137
Db	4131	CTGGCCCCGGCATCCAGCCGCGAGAAAGAGTCTTCAACTCCAGAGAAATTTAGTCGGCGT	4190
QY	4138	CTTAAGAAACGATATCACACAAATATTCCTCACCGATTTCAAGTATGACCTGAACATGCCA	4197
Db	4191	CTTAAGAAACGATATCACACAAATATTCCTCACCGATTTCAAGTATGACCTGAACATGCCA	4250
QY	4198	GCCAAAGTGTGCTGTGTGTCTGGAATACCGTGCACTTTTGAACGCCAGGCATCCAATGT	4257
Db	4251	GCCAAAGTGTGTGTGTGTCTGGAATACCGTGCACTTTTGAACGCCAGGCATCCAATGT	4310
QY	4258	CTCGAATGTCAAGTATGTGTCAACCCCAAGTCTCCAGTCTTGCACGCCACTTCGCGC	4317
Db	4311	CTCGAATGTCAAGTATGTGTCAACCCCAAGTCTCCAGTCTTGCACGCCACTTCGCGC	4370
QY	4318	TTTGCTGTGGAATATGCGACACACTTCACCGAGGCTTTGCGGTACAAATGAATCC	4377
Db	4371	TTGCTGTGGAATATGCGACACACTTCACCGAGGCTTTGCGGTACAAATGAATCC	4430
QY	4378	CCAGGTCTCCAGACCAAGAGGCCACAGACACTTGCACTTGGAAGGTTGATGAAGTGTG	4437
Db	4431	CCAGGTCTCCAGACCAAGAGGCCACAGACACTTGCACTTGGAAGGTTGATGAAGTGTG	4490
QY	4438	CCCAAGAAATTAACAAAGAGGCTGGGACAGGAAGTACATTTGTCTTGAGGGA	4497
Db	4491	CCCAAGAAATTAACAAAGAGGCTGGGACAGGAAGTACATTTGTCTTGAGGGA	4550
QY	4498	TCAAAAGTCTCATTTATATGACATGAAGCCACAGAAAGCTTGACAGAGGCCGTGGAAGAA	4557
Db	4551	TCAAAAGTCTCATTTATATGACATGAAGCCACAGAAAGCTTGACAGAGGCCGTGGAAGAA	4610
QY	4558	TTTGAAGTGTGGCTTCCCGACCGGGATGTATCTATTCAATGATGCCGTTGGTCTTCGAA	4617
Db	4611	TTTGAAGTGTGGCTTCCCGACCGGGATGTATCTATTCAATGATGCCGTTGGTCTTCGAA	4670
QY	4618	CTCGCAAAATACAGCCAAAGACAGATGTCCCATACATCTGAAGATGGAATCTCACCCGCAC	4677
Db	4671	CTCGCAAAATACAGCCAAAGACAGATGTCCCATACATCTGAAGATGGAATCTCACCCGCAC	4730
QY	4678	ACCACCTGTGGCCCGGGAGAAACCTCTTACTTGTGCTTCCAGCTTCCCTGCAAAACAG	4737
Db	4731	ACCACCTGTGGCCCGGGAGAAACCTCTTACTTGTGCTTCCAGCTTCCCTGCAAAACAG	4790
QY	4738	CGCTGGGTCAACGGCTTGAATCAGTGTGCGAGGTGGAGAGTTCTAGGGAAAAAGCA	4797
Db	4791	CGCTGGGTCAACGGCTTGAATCAGTGTGCGAGGTGGAGAGTTCTAGGGAAAAAGCA	4850
QY	4798	GAAGCTGATGTCTAACTGCTTGAAACTCCCTGTGAAACTGGAGAGTGTGACCGTCTA	4857
Db	4851	GAAGCTGATGTCTAACTGCTTGAAACTCCCTGTGAAACTGGAGAGTGTGACCGTCTA	4910
QY	4858	GAATGAACCTGACGCTGCCCTTCAATGACAGTGTGTGGGCAACCGAGAAAGG	4917
Db	4911	GAATGAACCTGACGCTGCCCTTCAATGACAGTGTGTGGGCAACCGAGAAAGG	4970
QY	4918	CTCTACGCCCTGAAGTCTTGA AAAAATCCCTTAACCCAGTCCCAAGAAATTTGAGGAGTC	4977
Db	4971	CTCTACGCCCTGAAGTCTTGA AAAAATCCCTTAACCCAGTCCCAAGAAATTTGAGGAGTC	5030
QY	4978	TTCCAAATTTATATTATCAAGAGACCTGGAAGCTACTCATGTATACAGAGAGAGACGG	5037
Db	5031	TTCCAAATTTATATTATCAAGAGACCTGGAAGCTACTCATGTATACAGAGAGAGACGG	5090
QY	5038	GCACTGTGTCTTTGTGACGTGAAGAAAGTGAACAGTCCCTGGCCCAATGCCACTGTGCT	5097
Db	5091	GCACTGTGTCTTTGTGACGTGAAGAAAGTGAACAGTCCCTGGCCCAATGCCACTGTGCT	5150
QY	5098	GCCCAAGCCGCAATCTCACCCCAATTTTTTGAAGCTGTCAAGGCGTGGCACTTGTTTGG	5157
Db	5151	GCCCAAGCCGCAATCTCACCCCAATTTTTTGAAGCTGTCAAGGCGTGGCACTTGTTTGG	5210

```

Qy 5158 GCAGGCAAGATTGAGAACTGGCTCTGCACTCTGTGAGCCATCCAGCAAGTGTGATT 5217
Db 5211 GCAGGCAAGATTGAGAACTGGCTCTGCACTCTGTGAGCCATCCAGCAAGTGTGATT 5270
Qy 5218 CTCGGCTACAAAGAAAAGCTCAGCAAACTACTGATCCGGAAGAGATAGAGACTCAGAG 5277
Db 5271 CTCGGCTACAAAGAAAAGCTCAGCAAACTACTGATCCGGAAGAGATAGAGACTCAGAG 5330
Qy 5278 CCTGCGAGCTGTATCCACTTCAACCAATTAAGTATCTCTATTGGAACCAATTAATTTCTAC 5337
Db 5331 CCTGCGAGCTGTATCCACTTCAACCAATTAAGTATCTCTATTGGAACCAATTAATTTCTAC 5390
Qy 5338 GAAATGACATGAAAGAGTACAGCTCAGAGAAATTCCTGGAATAGAAATGACCAATTCCTTG 5397
Db 5391 GAAATGACATGAAAGAGTACAGCTCAGAGAAATTCCTGGAATAGAAATGACCAATTCCTTG 5450
Qy 5398 GCACTGCTGTGTGTTGCGCTCTTCCAAAGACTTCCCTGTCTCAATCGTGAGGTGAAAC 5457
Db 5451 GCACTGCTGTGTGTTGCGCTCTTCCAAAGACTTCCCTGTCTCAATCGTGAGGTGAAAC 5510
Qy 5458 AGCGCAGGGGAGGAGAGAGATCTGTGTGTTTCCAAATTTGAGTGTGTTGAT 5517
Db 5511 AGCGCAGGGGAGGAGAGATCTGTGTGTTTCCAAATTTGAGTGTGTTGAT 5570
Qy 5518 TCTTACGGAAGAGATGAGCGGAGAGAGATCTGATGAGTGTGTTTCCAAATTTGAGTGTGAT 5577
Db 5571 TCTTACGGAAGAGATGAGCGGAGAGAGATCTGATGAGTGTGTTTCCAAATTTGAGTGTGAT 5630
Qy 5578 GCTTACAGAGAACCTTATCTGTTTGTGATCCCACTTCAACTCACTCGAATGATGATGATC 5637
Db 5631 GCTTACAGAGAACCTTATCTGTTTGTGATCCCACTTCAACTCACTCGAATGATGATGATC 5690
Qy 5638 CAGGCAAGCTCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5697
Db 5691 CAGGCAAGCTCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5750
Qy 5698 TACCTGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5757
Db 5751 TACCTGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5810
Qy 5758 TTAAGGATCATTTTGTCTGCAAGAGAGAACTGTGTAAGAGAGTCCCGCACTGAACCAACCG 5817
Db 5811 TTAAGGATCATTTTGTCTGCAAGAGAGAACTGTGTAAGAGAGTCCCGCACTGAACCAACCG 5870
Qy 5818 GGGCCGTCTCACTCTCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5877
Db 5871 GGGCCGTCTCACTCTCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5930
Qy 5878 ATCACCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5937
Db 5931 ATCACCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5990
Qy 5938 CCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5997
Db 5991 CCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6050
Qy 5998 GGGCCGCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6057
Db 6051 GGGCCGCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6110
Qy 6058 TCCCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6117
Db 6111 TCCCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6170
Qy 6118 ACCCGGCTGTCTCCAGAGTGAACAAAGTCTGGAGACAGTCTTCAAGTATATA 6165
Db 6171 ACCCGGCTGTCTCCAGAGTGAACAAAGTCTGGAGACAGTCTTCAAGTATATA 6218

```

RESULT 6
 AAL55217
 ID AAL55217 standard; DNA; 6156 BP.
 XX

```

AC AAL55217;
XX
DT 01-MAY-2003 (first entry)
XX
DE Human CR1K related DNA sequence, SEQ ID No 8.
XX
KW Anorectic; hypotensive; cardiast; antilipemic; cerebroprotective;
KW antitumor; osteopathic; antidiabetic; cytosolic; antidepressant;
KW immunomodulator; antitumor; tranquilizer; antiparkinsonian; nootropic;
KW neuroprotective; antidiabetic; antidiabetic; analgesic;
KW human citron rho/rac-interacting kinase; enzyme; CR1K; ameliorating;
KW obesity; comorbidity; cancer; anorexia; cachexia; bulimia;
KW central nervous system disorder; chronic obstructive pulmonary disease;
KW diabetes; pain; ds.
XX
OS Homo sapiens.
XX
PN WC02003004523-A1.
XX
PD 16-JAN-2003.
XX
PF 28-JUN-2002; 2002MO-EP07156.
XX
PR 02-JUL-2001; 2001JUS-0301841P.
PR 11-DEC-2001; 2001JUS-0338651P.
PR 25-APR-2002; 2002JUS-0375014P.
XX
PA (FARB ) BAYER AG.
XX
PI Zhu Z;
XX
DR WPI; 2003-221576/21.
XX
PT New human citron rho/rac-interacting kinase (CR1K) polypeptide and
PT polynucleotide, useful in preventing, ameliorating or treating diseases
PT associated with human CR1K dysfunction, e.g. obesity, diabetes or
PT Alzheimer's disease.
XX
PS Disclosure; Page 217-222; 237pp; English.
XX
CC The invention relates to an isolated polynucleotide encoding a human
CC citron rho/rac-interacting kinase polypeptide. The isolated
CC polynucleotide comprises a 6165 or 6603 base pair sequence, given in the
CC specification. The human citron rho/rac-interacting kinase (CR1K)
CC polypeptide and polynucleotide are useful in preventing, ameliorating, or
CC treating diseases associated with human CR1K dysfunction such as obesity
CC and obesity-associated comorbidities (e.g. hypertension, coronary artery
CC disease, hyperlipidaemia, stroke, gout, osteoarthritis, some types of
CC cancer including endometrial, breast, prostate and colon cancer),
CC anorexia, cachexia, bulimia, central nervous system disorders (e.g. mood
CC disorders, anxiety disorders, Parkinson's disease or Alzheimer's
CC disease), chronic obstructive pulmonary disease, or diabetes. These can
CC also be used to treat pain associated with the disorders. The human CR1K
CC polypeptide is also useful in diagnostic assays or in genetic testing.
CC The expression vector or the reagent is useful in preparing a medicament
CC for modulating the activity of a human CR1K in a disease, e.g. obesity, a
CC central nervous system disorder, or chronic obstructive pulmonary
CC disease. The fusion protein is useful for generating antibodies against a
CC CR1K polypeptide and for use in various assay systems. The methods are
CC useful in producing and detecting the polynucleotide and polypeptide and
CC in screening for agents that modulate the activity of the human CR1K
CC polypeptide. This polynucleotide sequence represents a DNA sequence
CC relating to the human CR1K protein of the invention
XX
SQ Sequence 6156 BP; 1732 A; 1548 C; 1679 G; 1197 T; 0 U; 0 Other;

```

Query Match 99.6%; Score 6142.2; DB 9; Length 6156;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 6147; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

```

Qy 1 ATGTTGAAGTCAATATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 60
Db 1 ATGTTGAAGTCAATATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 60

```


QY 61 GCCAGCCGGGCTTCAGGCTGAATCTGTCTTCCAGGGGAAACACCCTTTATGACTCAA 120
DB 61 GCCAGCCGGGCTTCAGGCTGAATCTGTCTTCCAGGGGAAACACCCTTTATGACTCAA 120
QY 121 CAGAGATGTCCTCTCTTCCCGAAGGGAATTAATGACCCCTTTGTTCTCTTGA 180
DB 121 CAGAGATGTCCTCTCTTCCCGAAGGGAATTAATGACCCCTTTGTTCTCTTGA 180
QY 181 GAATGCAAGCTGCTGTCTGATGAAATTAAAGCAGTGAACAATTGTCCGGAATAT 240
DB 181 GAATGCAAGCTGCTGTCTGATGAAATTAAAGCAGTGAACAATTGTCCGGAATAT 240
QY 241 TCCGACACCAATGCTGATTACAGAGCTCCAGCTTCGCGAAAGACTTCGAAGTGA 300
DB 241 TCCGACACCAATGCTGATTACAGAGCTCCAGCTTCGCGAAAGACTTCGAAGTGA 300
QY 301 AGCTCTGAGTGTGTGCTCACTTTCGAAAGTGAAGTGTGAAGAAAGCAACCGGG 360
DB 301 AGCTCTGAGTGTGTGCTCACTTTCGAAAGTGAAGTGTGAAGAAAGCAACCGGG 360
QY 361 GACATCTATCTATGAAGAATGATGAAGAAGGCTTTATTTGGCCAGAGAGTTTCA 420
DB 361 GACATCTATCTATGAAGAATGATGAAGAAGGCTTTATTTGGCCAGAGAGTTTCA 420
QY 421 TTTTGTGAGGAAGCGGAACATATTTATCTGAAGCAAGCCGCTGGATCCCCAATTA 480
DB 421 TTTTGTGAGGAAGCGGAACATATTTATCTGAAGCAAGCCGCTGGATCCCCAATTA 480
QY 481 CAGTATGCTCTTTCAGGACAAAATTCACCTTTATCTGGTCAATGAAATCAGCTGAGGG 540
DB 481 CAGTATGCTCTTTCAGGACAAAATTCACCTTTATCTGGTCAATGAAATCAGCTGAGGG 540
QY 541 GACTGCTGTCACTTTGAATGATATGAGGACAGTTAGATGAAACCTGATACAGTTT 600
DB 541 GACTGCTGTCACTTTGAATGATATGAGGACAGTTAGATGAAACCTGATACAGTTT 600
QY 601 TACCTAGCTAGCTGATTTTGGCTGTTACAGCGTTTCATCTGATGGGATACTGACATGA 660
DB 601 TACCTAGCTAGCTGATTTTGGCTGTTACAGCGTTTCATCTGATGGGATACTGACATGA 660
QY 661 GACATCAAGCTGAGAACTTCTGTTGACCGGACAGAGACATCAAGCTGTGTGATTTT 720
DB 661 GACATCAAGCTGAGAACTTCTGTTGACCGGACAGAGACATCAAGCTGTGTGATTTT 720
QY 721 GGATCTGCCGGAATAATGAATTCAAACAAGTGTGAATGCCAACTCCGATTTGGGACC 780
DB 721 GGATCTGCCGGAATAATGAATTCAAACAAGTGTGAATGCCAACTCCGATTTGGGACC 780
QY 781 CCAGATTACATGAGCTCCTGAAAGTCTGATGATGAAACGGGATGAAAAAGCACCTTAC 840
DB 781 CCAGATTACATGAGCTCCTGAAAGTCTGATGATGAAACGGGATGAAAAAGCACCTTAC 840
QY 841 GGGCTGGAAGCTGTGATGTGTGTGATGTGGGCTGATTTGCTATGAGATGATTAATGAGAGA 900
DB 841 GGGCTGGAAGCTGTGATGTGTGTGATGTGGGCTGATTTGCTATGAGATGATTAATGAGAGA 900
QY 901 TCCGCTCTGCAAGGGAACCTGTGCGCAAGACCTTCAATTAAGATTTTCCAGCGG 960
DB 901 TCCGCTCTGCAAGGGAACCTGTGCGCAAGACCTTCAATTAAGATTTTCCAGCGG 960
QY 961 TTTTGAATAATTTCCAGATGACCCCAAGTGAAGTGAATTTCTTGATCTGATTTCAAGC 1020
DB 961 TTTTGAATAATTTCCAGATGACCCCAAGTGAAGTGAATTTCTTGATCTGATTTCAAGC 1020
QY 1021 TTGTGTGCGGCGAAGAAAGAGACTGAAGTTGAAGTCTTTGCTGCTCCTTTCTTC 1080
DB 1021 TTGTGTGCGGCGAAGAAAGAGACTGAAGTTGAAGTCTTTGCTGCTCCTTTCTTC 1080
QY 1081 TCTAAAAATTGAATGGAACAATTCGTAATCTCTCTCCCTTCCTGTTCCACCTTCAG 1140
DB 1081 TCTAAAAATTGAATGGAACAATTCGTAATCTCTCTCCCTTCCTGTTCCACCTTCAG 1140

QY 1141 TCTGAGATGACACTTCCAAATTTTGAATGAACAGAGAAAGAAATTCGTGGGTTTCACTCCT 1200
DB 1141 TCTGAGATGACACTTCCAAATTTTGAATGAACAGAGAAAGAAATTCGTGGGTTTCACTCCT 1200
QY 1201 CCGTGCAGCTGAGCCCTCAGGCTTCTCGGGTGAAGAACTGCGGTTGTGGGGTTTTCG 1260
DB 1201 CCGTGCAGCTGAGCCCTCAGGCTTCTCGGGTGAAGAACTGCGGTTGTGGGGTTTTCG 1260
QY 1261 TACAGCAAGGCACTGGGGATTTTGGTGAATCTGAAGTCTGTGTGTGGGCTTGGACTCC 1320
DB 1261 TACAGCAAGGCACTGGGGATTTTGGTGAATCTGAAGTCTGTGTGTGGGCTTGGACTCC 1320
QY 1321 CTTGCCAAGACTGTCTTCATGAAAGAACTTCTCATCAAAAGCAAGGCTTCAAGAC 1380
DB 1321 CTTGCCAAGACTGTCTTCATGAAAGAACTTCTCATCAAAAGCAAGGCTTCAAGAC 1380
QY 1381 TCTCAGGACAAAGTGTACAAAGATGAGGAGGAAATGACCGGTTTACATGCGAGAGTGA 1440
DB 1381 TCTCAGGACAAAGTGTACAAAGATGAGGAGGAAATGACCGGTTTACATGCGAGAGTGA 1440
QY 1441 GAGTGAAGCTGTGCTTGTAGTCAGAAAGAGTGAAGCTGAAGGCTTCGAGACTCAGAGA 1500
DB 1441 GAGTGAAGCTGTGCTTGTAGTCAGAAAGAGTGAAGCTGAAGGCTTCGAGACTCAGAGA 1500
QY 1501 TCCCTCCTGAGGACGACCTTGTCTACCTACATCAAGATGCAATGCTTAAAGCGAAGT 1560
DB 1501 TCCCTCCTGAGGACGACCTTGTCTACCTACATCAAGATGCAATGCTTAAAGCGAAGT 1560
QY 1561 TTGGACCAAGCAAGATGAGGAGTGTCCAGAGAGATGACAAACCTGACGCTTCTCAT 1620
DB 1561 TTGGACCAAGCAAGATGAGGAGTGTCCAGAGAGATGACAAACCTGACGCTTCTCAT 1620
QY 1621 GATATCAGAGAGCAGAGCCGGAAGCTCCAAGAAATCAAGAGAGAGAGTACAGGCTCAA 1680
DB 1621 GATATCAGAGAGCAGAGCCGGAAGCTCCAAGAAATCAAGAGAGAGAGTACAGGCTCAA 1680
QY 1681 GTGGAAGAAATGAGGTTGATGAATCAAGTTGAAAGAGGATCTTGTCTCAGCAAGAGA 1740
DB 1681 GTGGAAGAAATGAGGTTGATGAATCAAGTTGAAAGAGGATCTTGTCTCAGCAAGAGA 1740
QY 1741 CGGAGTACTCTACAAATCTGAGCTGAGAGACTCTCGGCTTCTGTAAGAAATTCAG 1800
DB 1741 CGGAGTACTCTACAAATCTGAGCTGAGAGACTCTCGGCTTCTGTAAGAAATTCAG 1800
QY 1801 CGGAAAGCAGAGAAATGTCAGCATTAATCTGTTGAAGGCTTAAGATCAAGGAAAGCTGAA 1860
DB 1801 CGGAAAGCAGAGAAATGTCAGCATTAATCTGTTGAAGGCTTAAGATCAAGGAAAGCTGAA 1860
QY 1861 GTGGAGAAATATGCGAAATGAGAAATCAATGCTGAGCAGAGCTCAAAATTCAGAG 1920
DB 1861 GTGGAGAAATATGCGAAATGAGAAATCAATGCTGAGCAGAGCTCAAAATTCAGAG 1920
QY 1921 CTCGAAGAGAACTGAGAAAGGCTGTAAAGCCAGACCGAGGCTGAGGCTGAGGCTGAG 1980
DB 1921 CTCGAAGAGAACTGAGAAAGGCTGTAAAGCCAGACCGAGGCTGAGGCTGAGGCTGAG 1980
QY 1981 AATATCCGCGAGGCAAGGAGCGAGCGAGGAGGAGCTGGAAGAGCTGCAAGACCGAGAG 2040
DB 1981 AATATCCGCGAGGCAAGGAGCGAGCGAGGAGGAGCTGGAAGAGCTGCAAGACCGAGAG 2040
QY 2041 GATTTCTTGAAGGCTCAGAAAGAAAGCTGTGTGAAGCTGAAGGAGCGCGCATTTCTGT 2100
DB 2041 GATTTCTTGAAGGCTCAGAAAGAAAGCTGTGTGAAGCTGAAGGAGCGCGCATTTCTGT 2100
QY 2101 GAGAACAGGTTAAAGAGACTGAGACCATGAGACCTGAGAGAAACAGACTGAAGGATGAC 2160
DB 2101 GAGAACAGGTTAAAGAGACTGAGACCATGAGACCTGAGAGAAACAGACTGAAGGATGAC 2160
QY 2161 ATCCAGCAAAATCCCAACAGATCCAGCAGATGGCTGATTAATTTCTGAGCTCGAAGAG 2220
DB 2161 ATCCAGCAAAATCCCAACAGATCCAGCAGATGGCTGATTAATTTCTGAGCTCGAAGAG 2220
QY 2221 AAACATCGGAGGCGCAAGTCTCAGCCAGCACTTAAGAGTGAACCTGAAGCAAGAAAGAG 2280

Db 2221 AAACATGGAGAGGCCCAAGTCTCAGCCCAAGCACTTAAGTGCACCTGAAACAAAGAG 2280
 QY 2281 CAGCACTATGAGAAAAGATTAAAGTTGGACAATGATTAAGAAAGACTGGCTGAC 2340
 Db 2281 CAGCACTATGAGAAAAGATTAAAGTTGGACAATGATTAAGAAAGACTGGCTGAC 2340
 QY 2341 AAGGAGCACTGGAGAACATGATGACAGACACAGAGAGAGGCCCTGAGAGAGGCAAA 2400
 Db 2341 AAGGAGCACTGGAGAACATGATGACAGACACAGAGAGAGGCCCTGAGAGAGGCAAA 2400
 QY 2401 ATTCTGCGAAACGAGAGCGATGATCATGCTATGATTCAAAGATCAGATCCCTGGAA 2460
 Db 2401 ATTCTGCGAAACGAGAGCGATGATCATGCTATGATTCAAAGATCAGATCCCTGGAA 2460
 QY 2461 CAGAGATTGTGGAATGTTCTGTAAGCCAAATPAACTTGCAGCAATAGCAGTCTTTTACC 2520
 Db 2461 CAGAGATTGTGGAATGTTCTGTAAGCCAAATPAACTTGCAGCAATAGCAGTCTTTTACC 2520
 QY 2521 CAAGAGCACTGAAAGGCCCAAGAGATGATTTCTGAATCTCAGGCAACGAAATTTTAC 2580
 Db 2521 CAAGAGCACTGAAAGGCCCAAGAGATGATTTCTGAATCTCAGGCAACGAAATTTTAC 2580
 QY 2581 CTGAGACACAGGCTGGGAAATTGAGGCCCAGAACCGAAACCTGAGAGGAGCTGGAG 2640
 Db 2581 CTGAGACACAGGCTGGGAAATTGAGGCCCAGAACCGAAACCTGAGAGGAGCTGGAG 2640
 QY 2641 AAGATCAGCCACCAAGACCAAGTGAACAGATTCGGCTGTGGAACCTGAGACAAAGTTG 2700
 Db 2641 AAGATCAGCCACCAAGACCAAGTGAACAGATTCGGCTGTGGAACCTGAGACAAAGTTG 2700
 QY 2701 CCGGAGCTGAGTTAAGACACGAGAGCAGAACTGAGCTCAAGGCCCTCAAGAG 2760
 Db 2701 CCGGAGCTGAGTTAAGACACGAGAGCAGAACTGAGCTCAAGGCCCTCAAGAG 2760
 QY 2761 CTACAGCTCTCCCTGACGAGGCGGAGTCAAGTTGACAGCCCTGAGAGCTGACCGGCG 2820
 Db 2761 CTACAGCTCTCCCTGAGAGGCGGAGTCAAGTTGACAGCCCTGAGAGCTGACCGGCG 2820
 QY 2821 GCCCTGAGAGGCGGAGTTCGCAAGGCGAAGACAGAGCTGAGAGAGACACAGAGAGCT 2880
 Db 2821 GCCCTGAGAGGCGGAGTTCGCAAGGCGAAGACAGAGCTGAGAGAGACACAGAGAGCT 2880
 QY 2881 GAAAGAGAGATTCAGGCACTCAGGCACTAGAGATGAATCCAGGCAAAATTTGATGCT 2940
 Db 2881 GAAAGAGAGATTCAGGCACTCAGGCACTAGAGATGAATCCAGGCAAAATTTGATGCT 2940
 QY 2941 CTTGCTAAGAGCTGATCTGTAATCAGACCTGGAAGAGAGCTAAACAGCTGACCGAG 3000
 Db 2941 CTTGCTAAGAGCTGATCTGTAATCAGACCTGGAAGAGAGCTAAACAGCTGACCGAG 3000
 QY 3001 GACAAAGCTGAATCTCAACACCAAACTTCTATCTTCCAAACAACTCGATGAGCTTCT 3060
 Db 3001 GACAAAGCTGAATCTCAACACCAAACTTCTATCTTCCAAACAACTCGATGAGCTTCT 3060
 QY 3061 GGGCGCAACGAGAGATTGTACACTGCAAGTGAAGTGAACCATCTCCGCGGAGATC 3120
 Db 3061 GGGCGCAACGAGAGATTGTACACTGCAAGTGAAGTGAACCATCTCCGCGGAGATC 3120
 QY 3121 ACGGAACGAGATGAGCTTACAGCCAGAGCAACGATGAGGCTTGAAGACCAAG 3180
 Db 3121 ACGGAACGAGATGAGCTTACAGCCAGAGCAACGATGAGGCTTGAAGACCAAG 3180
 QY 3181 TGACCATGCTGAGAGAACGATGATTTGAGAGCCCTTAAACGATGAGCTGCTAGAA 3240
 Db 3181 TGACCATGCTGAGAGAACGATGATTTGAGAGCCCTTAAACGATGAGCTGCTAGAA 3240
 QY 3241 AAAAGCGGCACTGAGAGGCTTGAAGAGCGTCTTGAGTGAAGAAATCCAGTTTGAAG 3300
 Db 3241 AAAAGCGGCACTGAGAGGCTTGAAGAGCGTCTTGAGTGAAGAAATCCAGTTTGAAG 3300
 QY 3301 TGTCCGGTTCCGAGAGCTGACAGATGCTGACACCGAGAAACACAGAGGCGGAGAGCC 3360
 Db 3301 TGTCCGGTTCCGAGAGCTGACAGATGCTGACACCGAGAAACACAGAGGCGGAGAGCC 3360

Db 3301 TGTCCGGTTCCGAGAGCTGACAGATGCTGACACCGAGAAACACAGAGGCGGAGAGCC 3360
 QY 3361 GATCAGCGGATCAACGAGTCTGCGCAAGTGTGAGAGCTGCGACATGAAAGAGCAACAGCT 3420
 Db 3361 GATCAGCGGATCAACGAGTCTGCGCAAGTGTGAGAGCTGCGACATGAAAGAGCAACAGCT 3420
 QY 3421 GAGATTCTCGCTCTGACAGAGCTTCTCAAGAGCAAGAGCTGAAGGCCGAGAGCTTCT 3480
 Db 3421 GAGATTCTCGCTCTGACAGAGCTTCTCAAGAGCAAGAGCTGAAGGCCGAGAGCTTCT 3480
 QY 3481 GACAGCTCAATGACCTGAGAGAGAGCAATGCTATGCTTGAATGAAATGCGGAAAGTTA 3540
 Db 3481 GACAGCTCAATGACCTGAGAGAGAGCAATGCTATGCTTGAATGAAATGCGGAAAGTTA 3540
 QY 3541 CAGCAGAGCTGAGAGCTGAACGAGAGCTCAACAGAGGCTTCTGGAAGAGCAAGCCAAA 3600
 Db 3541 CAGCAGAGCTGAGAGCTGAACGAGAGGCTCAACAGAGGCTTCTGGAAGAGCAAGCCAAA 3600
 QY 3601 TTTACAGCAGATGGAAGCTGCAAGAAAATCAGATTTTCCGTGACTCAAGAGCTGCAA 3660
 Db 3601 TTTACAGCAGATGGAAGCTGCAAGAAAATCAGATTTTCCGTGACTCAAGAGCTGCAA 3660
 QY 3661 GAAAGCTTGAATGGGCTGATCTACTGAGAGACAGAAAGAGTGAAGCTTGAAGATTCAGCTG 3720
 Db 3661 GAAAGCTTGAATGGGCTGATCTACTGAGAGACAGAAAGAGTGAAGTGAAGATTCAGCTG 3720
 QY 3721 GAAACATTCAAGTTCTATTCTCATGAAAGGTGAAATGGAAGGCACTATTCTCAA 3780
 Db 3721 GAAACATTCAAGTTCTATTCTCATGAAAGGTGAAATGGAAGGCACTATTCTCAA 3780
 QY 3781 CAAACCAATCTGATGATTTTCTGCAAGCCAAATGGAACCTGCTTAAAAAGAAAAG 3840
 Db 3781 CAAACCAATCTGATGATTTTCTGCAAGCCAAATGGAACCTGCTTAAAAAGAAAAG 3840
 QY 3841 GTTCCCTGTCAGATGAATGAGCTGGAAGCTGGGCCCTGGAAGAGAAAGCTGCTGTGA 3900
 Db 3841 GTTCCCTGTCAGATGAATGAGCTGGAAGCTGGGCCCTGGAAGAGAAAGCTGCTGTGA 3900
 QY 3901 GAGCTAAGAGAGCCCTTTCAGAAAGCCCGCATCGAGCTCCGCTCCGCGGAGAGAGCT 3960
 Db 3901 GAGCTAAGAGAGCCCTTTCAGAAAGCCCGCATCGAGCTCCGCTCCGCGGAGAGAGCT 3960
 QY 3961 GCCCACCGCAAGCAACGAGCAACCAACCAATCCACGCTCAGCCACCGGAGAGCAAG 4020
 Db 3961 GCCCACCGCAAGCAACGAGCAACCAACCAATCCACGCTCAGCCACCGGAGAGCAAG 4020
 QY 4021 ATGCGCATGTCGCGCATGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 4080
 Db 4021 ATGCGCATGTCGCGCATGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 4080
 QY 4081 GCCCGGCAATCCAGCGCGAGAAAGAGTCTTCAACTCCAGAGAAATTTAGTCGGGCTTT 4140
 Db 4081 GCCCGGCAATCCAGCGCGAGAAAGAGTCTTCAACTCCAGAGAAATTTAGTCGGGCTTT 4140
 QY 4141 AAGGAACGATGACCAACATATTCCTCAACGATTCAGCTGAGACTGAACATGCGAGCC 4200
 Db 4141 AAGGAACGATGACCAACATATTCCTCAACGATTCAGCTGAGACTGAACATGCGAGCC 4200
 QY 4201 ACAAAGTGTGTCGTGTGTGTAACCGTGCACTTTGACGCGCAAGCAATCCAAATGTCTC 4260
 Db 4201 ACAAAGTGTGTCGTGTGTGTAACCGTGCACTTTGACGCGCAAGCAATCCAAATGTCTC 4260
 QY 4261 GAATGTCAAGTGTGTGTCAACCCCAAGTGTCAAGTGTCTTGCAGCCACTTGCAGGCTTG 4320
 Db 4261 GAATGTCAAGT 4320
 QY 4321 CCTGCTGAATATGCAACACTTCAACGAGGCTTCTCCGCTGACAAATGAAATCTCCCA 4380
 Db 4321 CCTGCTGAATATGCAACACTTCAACGAGGCTTCTCCGCTGACAAATGAAATCTCCCA 4380
 QY 4381 GGTCTCCAGACCAAGAGAGCCGAGAGCTTGGACCTTGGAAAGGTGTGAAGGTGTGCC 4440
 Db 4381 GGTCTCCAGACCAAGAGAGCCGAGAGCTTGGACCTTGGAAAGGTGTGAAGGTGTGCC 4440

QY 4441 AGAATAACAAAGAGACAGCAAGGCTGGACAGAAATGATCATTTGCTTGGAGGATCA 4500
 DB 4441 AGGAATTAACAAAGAGACAGCAAGGCTGGACAGAAATGATCATTTGCTTGGAGGATCA 4500
 QY 4501 AAAGTCTCTTTATATGACAAATGAAGCCAGAAAGCTGGACAGAGCCGGTGGAAATTT 4560
 DB 4501 AAAGTCTCTTTATATGACAAATGAAGCCAGAAAGCTGGACAGAGCCGGTGGAAATTT 4560
 QY 4561 GAGCTGTCCTTCCGACGGGGATGATCTATTCAATGATGTCCTGCTTCCGAACTC 4620
 DB 4561 GAGCTGTCCTTCCGACGGGGATGATCTATTCAATGATGTCCTGCTTCCGAACTC 4620
 QY 4621 GCAAAATACAGCCAAAGAGATGTCCTATACATCTGAAGATGAAATCTCACCCGACACC 4680
 DB 4621 GCAAAATACAGCCAAAGAGATGTCCTATACATCTGAAGATGAAATCTCACCCGACACC 4680
 QY 4681 ACCTGTCGCGCCGGGAGAAACCTCTACTTGTAGTCTCCAGCTTCCCTGACAAACAGCCG 4740
 DB 4681 ACCTGTCGCGCCGGGAGAAACCTCTACTTGTAGTCTCCAGCTTCCCTGACAAACAGCCG 4740
 QY 4741 TGGGTACCGCCTTAGAATGATGTCGAGTGGAGAGTTTCTAGGAAAAAGCAGAA 4800
 DB 4741 TGGGTACCGCCTTAGAATGATGTCGAGTGGAGAGTTTCTAGGAAAAAGCAGAA 4800
 QY 4801 GCTGATGCTAACTGCTTGGAACTCCCTGCTGAACTGGAAGGTGATGACCTGTAGAC 4860
 DB 4801 GCTGATGCTAACTGCTTGGAACTCCCTGCTGAACTGGAAGGTGATGACCTGTAGAC 4860
 QY 4861 ATGAATGTCACGCTGCTTCACTGATGACCAAGTGTGTTGGTGGCACCCGAGAGGGCTC 4920
 DB 4861 ATGAATGTCACGCTGCTTCACTGATGACCAAGTGTGTTGGTGGCACCCGAGAGGGCTC 4920
 QY 4921 TAGCCCTGATGCTTGAATACTCCCTAACCAGATGCTCCAGGAAATTTGAGAGAGTTC 4980
 DB 4921 TAGCCCTGATGCTTGAATACTCCCTAACCAGATGCTCCAGGAAATTTGAGAGAGTTC 4980
 QY 4981 CAAATTTATATTAATCAAGACCTGGAAGCTACTCATATGATAGCAGAGAAAGCCGCA 5040
 DB 4981 CAAATTTATATTAATCAAGACCTGGAAGCTACTCATATGATAGCAGAGAAAGCCGCA 5040
 QY 5041 CTGTGCTTGTGACGTTGAAGAAAGTAAACAGTCCCTGGCCCACTGCTGCC 5100
 DB 5041 CTGTGCTTGTGACGTTGAAGAAAGTAAACAGTCCCTGGCCCACTGCTGCC 5100
 QY 5101 CAGCCCAATCTCAACCAATTTTGAAGCTGTCAAGGGCTGCCACTTGTGGGGCA 5160
 DB 5101 CAGCCCAATCTCAACCAATTTTGAAGCTGTCAAGGGCTGCCACTTGTGGGGCA 5160
 QY 5161 GGCAGATTGAGAACGGGCTCTGATCTGTGACAGCCATGCCAGCAAAATGCTCATTTCTC 5220
 DB 5161 GGCAGATTGAGAACGGGCTCTGATCTGTGACAGCCATGCCAGCAAAATGCTCATTTCTC 5220
 QY 5221 CGCTACACAGAAAACCTCAGCAATTAATCTGATCCGGAAGAGATGAGACCTCAGACCC 5280
 DB 5221 CGCTACACAGAAAACCTCAGCAATTAATCTGATCCGGAAGAGATGAGACCTCAGACCC 5280
 QY 5281 TGCAGCTGTATCCACTTCACTCAATTAAGTATCTCATTTGGAACCAATTAATTTAGGAA 5340
 DB 5281 TGCAGCTGTATCCACTTCACTCAATTAAGTATCTCATTTGGAACCAATTAATTTAGGAA 5340
 QY 5341 ATGCAATGAAGAGTACAGCTGAGAAATTCCTGATAGAAATGACATTCCTTGGCA 5400
 DB 5341 ATGCAATGAAGAGTACAGCTGAGAAATTCCTGATAGAAATGACATTCCTTGGCA 5400
 QY 5401 CTTGCTGTGTTTGGCGCTCTTCCACAGCTTCCCTGTCTCAATGTCGACAGTGAACAGC 5460
 DB 5401 CTTGCTGTGTTTGGCGCTCTTCCACAGCTTCCCTGTCTCAATGTCGACAGTGAACAGC 5460
 QY 5461 GCAGGGCAGGAGAGTACTTGTGTTTCCACGAATTTGAGAGTTCGTGGATTC 5520
 DB 5461 GCAGGGCAGGAGAGTACTTGTGTTTCCACGAATTTGAGAGTTCGTGGATTC 5520

QY 5521 TACGGAAGAGTACCGGACAGACGATCTCAATGAGTGTGCTTACCTTTGGCTTTGCC 5580
 DB 5521 TACGGAAGAGTACCGGACAGACGATCTCAATGAGTGTGCTTACCTTTGGCTTTGCC 5580
 QY 5581 TACAGGAACCCATCTGTTTGTGACCCACTTCAACTCACTGAAATTAATGATCAG 5640
 DB 5581 TACAGGAACCCATCTGTTTGTGACCCACTTCAACTCACTGAAATTAATGATCAG 5640
 QY 5641 GCAAGCTCTCAGCAGGAGCCCTGCGGACGTACTGTGACATCCGGAACCCGCGCTAC 5700
 DB 5641 GCAAGCTCTCAGCAGGAGCCCTGCGGACGTACTGTGACATCCGGAACCCGCGCTAC 5700
 QY 5701 CTGGGCTCTGCACTTTCCTTCAAGAGCATTTACTTGGCTGCTCATACAGATTAATTA 5760
 DB 5701 CTGGGCTCTGCACTTTCCTTCAAGAGCATTTACTTGGCTGCTCATACAGATTAATTA 5760
 QY 5761 AGGGTCAATTTGCTGCAAGGGGAACTCTGTAAGAGTCCGGCACTGAACACACCGGGGC 5820
 DB 5761 AGGGTCAATTTGCTGCAAGGGGAACTCTGTAAGAGTCCGGCACTGAACACACCGGGGC 5820
 QY 5821 CCGTCCACTCCCGCAGCAGCCCAACAGCAGGCCCCCACTGACAACTGACATC 5880
 DB 5821 CCGTCCACTCCCGCAGCAGCCCAACAGCAGGCCCCCACTGACAACTGACATC 5880
 QY 5881 ACCAAGCGCTGCTCTTCAAGCCAGCGCGCCGGAAGCCCCAGCCACCCGCGAGGCCA 5940
 DB 5881 ACCAAGCGCTGCTCTTCAAGCCAGCGCGCCGGAAGCCCCAGCCACCCGCGAGGCCA 5940
 QY 5941 AGGACACCCCAACGCTACCGCGAGGGCGGACCGAGCTGGCGAGGACAAGTCTCTGGC 6000
 DB 5941 AGGACACCCCAACGCTACCGCGAGGGCGGACCGAGCTGGCGAGGACAAGTCTCTGGC 6000
 QY 6001 CGCCCTGAGGAGAGAGAGTCCCCCGGCGGATACTGACACGCGAGAGACGGTCC 6060
 DB 6001 CGCCCTGAGGAGAGAGAGTCCCCCGGCGGATACTGACACGCGAGAGAGCGGTCC 6060
 QY 6061 CCCGCAAGCTGTTTGAAGCAGCAGCAGGAGGCGGCTGCTCGGAGCGCTGAGGACC 6120
 DB 6061 CCCGCAAGCTGTTTGAAGCAGCAGCAGGAGGCGGCTGCTCGGAGCGCTGAGGACC 6120
 QY 6121 CCGCTGTCCAGGTGAACAAAGTCTGGGACCACTC 6155
 DB 6121 CCGCTGTCCAGGTGAACAAAGTCTGGGACCACTC 6155

RESULT 7
 AAS06701 ID AAS06701 standard; cDNA; 6159 BP.
 XX AAS06701;
 DT 12-SEP-2001 (first entry)
 XX
 DE Polynucleotide sequence encoding human protein kinase #1.
 XX
 KW Human; protein kinase; PTK; STK; cancer; cardiovascular disease;
 KW metabolic disorder; immune related disease; neurological disorder;
 KW neurodegenerative disorder; inflammatory disorder; infectious disease;
 KW reproductive disorder; gene therapy; ss.
 XX
 OS Homo sapiens.
 XX
 PN W0200138503-A2.
 XX
 PD 31-MAY-2001.
 XX
 XX 22-NOV-2000; 2000MO-US032085.
 PF
 PR 24-NOV-1999; 99US-0167482P.
 XX
 XX (SUGB-) SUGEN INC.
 PA
 XX Plowman GD, Whyte D, Manning G, Sudarsanam S, Martinez R;

PI Planagan P, Clary D;
XX WPI: 2001-343950/36.
DR P-PSDB; AA003501.
XX Nucleic acids encoding human kinase polypeptides, useful for preventing
PT diagnosing and/or treating e.g. cancer, immune, cardiovascular and
PT neuronal-associated diseases, and microbial infections.
XX
PS Example 1; Fig 1; 433bp; English.
XX
CC AA06701-AA06757 encode for novel human protein kinases #1-57. The novel
CC protein kinases have been identified as members of the tyrosine or
CC serine/threonine kinase (PTK and STK) families. The polynucleotides
CC encoding protein kinases and the polypeptides may be used in the
CC inactivation, diagnosis and treatment of diseases associated with
CC inappropriate kinase expression. For example, they may be used to treat
CC cancers (especially cancers of hematopoietic origin), cardiovascular
CC disease (e.g. atherosclerosis), metabolic disorders (e.g. diabetes),
CC immune related diseases (e.g. rheumatoid arthritis), neurological
CC disorders (e.g. schizophrenia), neurodegenerative disorders (e.g.
CC Parkinson's disease), inflammatory disorders (e.g. asthma), infectious
CC disease (e.g. HIV) and reproductive disorders (e.g. infertility).
CC Additionally, polynucleotides encoding protein kinases may be used for
CC gene therapy and as DNA probes in diagnostic assays. The protein kinase
CC polypeptides may be used as antigens in the production of antibodies
CC against the protein kinases and in assays to identify modulators of
CC protein kinase expression and activity
CC
SQ Sequence 6159 BP; 1732 A; 1549 C; 1680 G; 1198 T; 0 U; 0 Other;

Query Match 99.4%; Score 6126; DB 4; Length 6159;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 6145; Conservative 0; Mismatches 10; Indels 3; Gaps 1;

QY 1 ATGTGAAGTTCAAAATATGAGCGCGGAATCTTTGGATGCTGTGCTGAACCAT 60
DB 1 ATGTGAAGTTCAAAATATGAGCGCGGAATCTTTGGATGCTGTGCTGAACCAT 60

QY 61 GCCAGCGGGCTCCAGGCTGAATCTGTTCTCCAGGGGAAACACCTTTATGACTCAA 120
DB 61 GCCAGCGGGCTCCAGGCTGAATCTGTTCTCCAGGGGAAACACCTTTATGACTCAA 120

QY 121 CAGCAATGCTCTCTCTTCCGAGAAAGGATTTAGATGCTCTTGTCTCTTGA 180
DB 121 CAGCAATGCTCTCTCTTCCGAGAAAGGATTTAGATGCTCTTGTCTCTTGA 180

QY 121 CAGCAATGCTCTCTCTTCCGAGAAAGGATTTAGATGCTCTTGTCTCTTGA 180
DB 121 CAGCAATGCTCTCTCTTCCGAGAAAGGATTTAGATGCTCTTGTCTCTTGA 180

QY 181 GAATGAGTCAGGCTGCTGTATGAGATTAAGCAGTGAGCACTTTGCCGGAA--G 237
DB 181 GAATGAGTCAGGCTGCTGTATGAGATTAAGCAGTGAGCACTTTGCCGGAA--G 240

QY 238 TATTCGACACCATAGCTAGTTACAGAGCTCCAGCTTCGCAAAAGACTTCGAAGTC 297
DB 241 TATTCGACACCATAGCTAGTTACAGAGCTCCAGCTTCGCAAAAGACTTCGAAGTC 300

QY 298 AGAAGCTCTTGAGGTTGTGCTCACTTGTCTGAAGTGCAGGTGTAAAGAAAGCAACC 357
DB 301 AGAAGCTCTTGAGGTTGTGCTCACTTGTCTGAAGTGCAGGTGTAAAGAAAGCAACC 360

QY 358 GGGGACATCTATGCTATGAAAGTGAAGAAAGAGGCTTTATTTGGCCAGAGAGGTT 417
DB 361 GGGGACATCTATGCTATGAAAGTGAAGAAAGAGGCTTTATTTGGCCAGAGAGGTT 420

QY 418 TCATTTTTGAGAGAGCGGAACATATTTATCTCGAAGCACAAGCCGCGATCCCCAA 477
DB 421 TCATTTTTGAGAGAGCGGAACATATTTATCTCGAAGCACAAGCCGCGATCCCCAA 480

QY 478 TTACAGTATGCTTTTCAGAGCAAAATCACTTTATCTGCTCATGGAATATCAGCTGGA 537
DB 481 TTACAGTATGCTTTTCAGAGCAAAATCACTTTATCTGCTCATGGAATATCAGCTGGA 540

QY 538 GGGGACTTGTCTGCTTTTGAATGATATGAGCACTTATGATGAAAACCTGATACG 597
DB 541 GGGGACTTGTCTGCTTTTGAATGATATGAGCACTTATGATGAAAACCTGATACG 600

DB 541 GGGGACTTGTCTGCTTTTGAATGATATGAGCACTTATGATGAAAACCTGATACG 600
QY 598 TTTTACCTAGCTAGCTGATTTTGGCTGTTCAACAGCGTTCATGTATGAGTATGTCAT 657
DB 601 TTTTACCTAGCTAGCTGATTTTGGCTGTTCAACAGCGTTCATGTATGAGTATGTCAT 660
QY 658 CGAGATTCAGAGCTTGAGAACTTTCTGTTGACCCGACAGACATCATAGCTGTGTGAT 717
DB 661 CGAGATTCAGAGCTTGAGAACTTTCTGTTGACCCGACAGACATCATAGCTGTGTGAT 720
QY 718 TTTGATCTGCGCGAAATGATTTCAACAGATGGTGAATGCCAACTCCGATTTGG 777
DB 721 TTTGATCTGCGCGAAATGATTTCAACAGATGGTGAATGCCAACTCCGATTTGG 780
QY 778 ACCCAAGATTACATGAGCTCTGGAAGTCTGATCTGATGAACCGGGATGGAAGGACCC 837
DB 781 ACCCAAGATTACATGAGCTCTGGAAGTCTGATCTGATGAACCGGGATGGAAGGACCC 840
QY 838 TACGCTGAGCTGTGATCTGTGTGTCACTGTGTGTGTGTGTGTGTGTGTGTGTGT 897
DB 841 TACGCTGAGCTGTGATCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 900
QY 898 AGATCCCTCTGAGAGGGAACCTGTGACAGAACTTCAATTAATTAATTTCCAG 957
DB 901 AGATCCCTCTGAGAGGGAACCTGTGACAGAACTTCAATTAATTAATTTCCAG 960
QY 958 CGGTTTTGAATTTCCAGATGAACCCCAAGTGAAGCACTTCTTGTATGATTTCAA 1017
DB 961 CGGTTTTGAATTTCCAGATGAACCCCAAGTGAAGCACTTCTTGTATGATTTCAA 1020
QY 1018 AGCTTTGTGTGGCCAGAAAGAGACTGAAGTTGAAGTCTTGTGCTGATCTTTTC 1077
DB 1021 AGCTTTGTGTGGCCAGAAAGAGACTGAAGTTGAAGTCTTGTGCTGATCTTTTC 1080
QY 1078 TTCTTAAATTAATGATCTGAACCAATGTAATCTCTCCCTCCCTGCTTCCACCTTC 1137
DB 1081 TTCTTAAATTAATGATCTGAACCAATGTAATCTCTCTCCCTGCTTCCACCTTC 1140
QY 1138 AAGTTCAGATGACACTTCAATTTTGAATGAACAGAAAGAAATCTGTGGTTTCATCC 1197
DB 1141 AAGTTCAGATGACACTTCAATTTTGAATGAACAGAAAGAAATCTGTGGTTTCATCC 1200
QY 1198 TCTCGTGCAGCTGAGCCCTCAAGCTTCTGGGTGAAGAACTGCGTTTGTGGGTTT 1257
DB 1201 TCTCGTGCAGCTGAGCCCTCAAGCTTCTGGGTGAAGAACTGCGTTTGTGGGTTT 1260
QY 1258 TCGTACAGCAAGCACTGGGGATCTTGGTATGATGATGATGATGATGATGATGATG 1317
DB 1261 TCGTACAGCAAGCACTGGGGATCTTGGTATGATGATGATGATGATGATGATGATG 1320
QY 1318 TCCCTGCGCAAGCTAGCTTCAATGAAAGAACTTTCATCAAAAGCAAGAGCTACAA 1377
DB 1321 TCCCTGCGCAAGCTAGCTTCAATGAAAGAACTTTCATCAAAAGCAAGAGCTACAA 1380
QY 1378 GACTCTCAGGACAGTGTCACAAGTGAAGCAGAAATGACCCGGTTACATGAGAGTG 1437
DB 1381 GACTCTCAGGACAGTGTCACAAGTGAAGCAGAAATGACCCGGTTACATGAGAGTG 1440
QY 1438 TCAGAGGTGAGGCTGTGTGTGTGAAGAGAGGTGTGAAGGCTCTGAGACTGAG 1497
DB 1441 TCAGAGGTGAGGCTGTGTGTGTGAAGAGAGGTGTGAAGGCTCTGAGACTGAG 1500
QY 1498 AGATCCCTCTGAGACAGGACTTGTCTACATCAAGATGAGTGAAGTGAAGTGAAG 1557
DB 1501 AGATCCCTCTGAGACAGGACTTGTCTACATCAAGATGAGTGAAGTGAAGTGAAG 1560
QY 1558 AGTTTGAAGCAAGCAGATGAGGTGTCAGAGAGATGACAAAGCACTGAGCTTCTC 1617
DB 1561 AGTTTGAAGCAAGCAGATGAGGTGTCAGAGAGATGACAAAGCACTGAGCTTCTC 1620
QY 1618 CATGATATCAGAGACAGAGCCGGAAGCTCCAAAGAAATCAAGAGCAGAGATGACAGCT 1677
DB 1621 CATGATATCAGAGACAGAGCCGGAAGCTCCAAAGAAATCAAGAGCAGAGATGACAGCT 1680

Qy	1678	UAACTGGAGAAATAGAGTTGATATGATATCAATTTGGAAAGAGATCTTTGTCTCAGCAGA	1737
Dp	1681	CAAGTGGAAAGAAATGAGGGTTGATATGATATCAATTTGGAAAGAGATCTTTGTCTCAGCAGA	1740
Qy	1738	AGA CGAGTGAATCTCTACGATCTGAGCTGAGAGATCTCGGCTTGCTGCGTGAAGATTC	1797
Dp	1741	AGACGGAGTGAATCTCTACGATCTGAGCTGAGAGATCTCGGCTTGCTGCGTGAAGATTC	1800
Qy	1798	AAGCGAAAGCGAGCAAGATCTCAGCATTAACCTGTTGAAGCTTAAGATCAAGGAAAGCT	1857
Dp	1801	AAGCGAAAGCGAGCAAGATCTCAGCATTAACCTGTTGAAGCTTAAGATCAAGGAAAGCT	1860
Qy	1858	GAAGTGGGAGATTTTCGAAACTGAGAGAGATCAATGCTGACAGCAGCTCAAAATTCAG	1917
Dp	1861	GAAGTGGGAGATTTTCGAAACTGAGAGAGATCAATGCTGACAGCAGCTCAAAATTCAG	1920
Qy	1918	GAGCTCCAGAGAAACTGGAGAAAGCGCTGTAATAACCGACAGAGGCCACCGAAGCTCTG	1977
Dp	1921	GAGCTCCAGAGAAACTGGAGAAAGCGCTGTAATAACCGACAGAGGCCACCGAAGCTCTG	1980
Qy	1978	CAGAAATATCCCGCCAGGCAAGAGAGCGAGCCGAGAGGGGAGCTGAGAGAGCTSCAAGCCGA	2037
Dp	1981	CAGAAATATCCCGCCAGGCAAGAGAGCGAGCCGAGAGGGGAGCTGAGAGAGCTSCAAGCCGA	2040
Qy	2038	GAGGATTTCTTGTGAAGGCATCAGAAAGAGCTGTGGAGAGCTGAGAAACGCGCGCATTC	2097
Dp	2041	GAGGATTTCTTGTGAAGGCATCAGAAAGAGCTGTGGAGAGCTGAGAAACGCGCGCATTC	2100
Qy	2098	CTGGAGAAACAAGGTAAAGAGACTAGAGACCATGAGCGTAAAGAAAACAAGCTAAGAT	2157
Dp	2101	CTGGAGAAACAAGGTAAAGAGACTAGAGACCATGAGCGTAAAGAAAACAAGCTAAGAT	2160
Qy	2158	GACATCCAGACAAAATCCCAACAGTCCAGAGAGATGGCTGATAAATTTCTGGAAGCTGAA	2217
Dp	2161	GACATCCAGACAAAATCCCAACAGTCCAGAGAGATGGCTGATAAATTTCTGGAAGCTGAA	2220
Qy	2218	GAGAAACATCGGAGGCGCCAAAGTCTCAGCCAGACACTAGAAAGTCACTTGAACAGAAA	2277
Dp	2221	GAGAAACATCGGAGGCGCCAAAGTCTCAGCCAGACACTAGAAAGTCACTTGAACAGAAA	2280
Qy	2278	GAGCAGCACTATGAGGAAAGATTTAAAGTGTTCGACAAATCAATTAAGAAAGACTTGCT	2337
Dp	2281	GAGCAGCACTATGAGGAAAGATTTAAAGTGTTCGACAAATCAATTAAGAAAGACTTGCT	2340
Qy	2338	GACAAAGGAGACATCGAGAGAAATATGAGAGACAGAGAGAGGCGCCATGAAAGAGCG	2397
Dp	2341	GACAAAGGAGACATCGAGAGAAATATGAGAGACAGAGAGAGGCGCCATGAAAGAGCG	2400
Qy	2398	AAAAATTCACGCGAAACAGAGAGCGATGATCAATGCTTGAATTCGAAGTCAAGATCCCTG	2457
Dp	2401	AAAAATTCACGCGAAACAGAGAGCGATGATCAATGCTTGAATTCGAAGTCAAGATCCCTG	2460
Qy	2458	GAAACAGAGATTTGGAACTGTCTGAAGCCAAATAACTTGCAGCAAAATAGACGTCCTTTT	2517
Dp	2461	GAAACAGAGATTTGGAACTGTCTGAAGCCAAATAACTTGCAGCAAAATAGACGTCCTTTT	2520
Qy	2518	AACCCAAAGAACATGAAAGGCCCAAGAAAGATGATTTCTGAATCTCAGCGCAACAGAAATTT	2577
Dp	2521	AACCCAAAGAACATGAAAGGCCCAAGAAAGATGATTTCTGAATCTCAGCGCAACAGAAATTT	2580
Qy	2578	TACCTGGAGACACAGGCTGGAGAGTTGGAGGCCCAAGAACGAAAACTGGAGAGCAGCTG	2637
Dp	2581	TACCTGGAGACACAGGCTGGAGAGTTGGAGGCCCAAGAACGAAAACTGGAGAGCAGCTG	2640
Qy	2638	GAGAAAGATCAGCCACCAAGACCAACATGACAAAGATCGGCTGCTGGAATCTGGAGACAGA	2697
Dp	2641	GAGAAAGATCAGCCACCAAGACCAACATGACAAAGATCGGCTGCTGGAATCTGGAGACAGA	2700
Qy	2698	TTGGCGGAGGCTCAGTCTTAGAGACAGAGAGCAAGAACTGAGAGCTCAGAGCGCAAGCTCA	2757
Dp	2701	TTGGCGGAGGCTCAGTCTTAGAGACAGAGAGCAAGAACTGAGAGCTCAGAGCGCAAGCTCA	2760

OY	2758	GAGCTACAGCTTCCTCCCTGACAGAGCGGGAGTCAACGTTGACAGCCCTGACAGGCTGCACCG	2817
Db	2761	GAGCTACAGCTTCCTCCCTGACAGAGCGGGAGTCAACGTTGACAGCCCTGACAGGCTGCACCG	2820
OY	2818	GCGGCCCTGAGAGCCAGCTTCGCGCAGGCGAAGACAGAGCTGGAAGAGACCAACAGCAGAA	2877
Db	2821	GCGGCCCTGAGAGCCAGCTTCGCGCAGGCGAAGACAGAGCTGGAAGAGACCAACAGCAGAA	2880
OY	2878	GCTGAGAGAGAGATCCAGGCACTTCACGGCACATGAGATGAAATCCAGCGCAAAATTTGAT	2937
Db	2881	GCTGAGAGAGAGATCCAGGCACTTCACGGCACATGAGATGAAATCCAGCGCAAAATTTGAT	2940
OY	2938	GCTTTTGTATACAGCTGTACTGTATACAGACTGAGAGAGAGCTTAAACAGCTGAC	2997
Db	2941	GCTTTTGTATACAGCTGTACTGTATACAGACTGAGAGAGAGCTTAAACAGCTGAC	3000
OY	2998	GAGGACAAAGCTGAACTCAACCAACCAAACTTCTACTGTCGCAAAACAATCTGATGAGGCT	3057
Db	3001	GAGGACAAAGCTGAACTCAACCAACCAAACTTCTACTGTCGCAAAACAATCTGATGAGGCT	3060
OY	3058	TCGCGCCCAACGACGAGATTGTACAACCTGCGAAGTAGAGTGCACATCTCCGCGGGAG	3117
Db	3061	TCGCGCCCAACGACGAGATTGTACAACCTGCGAAGTAGAGTGCACATCTCCGCGGGAG	3120
OY	3118	ATCACGGAACGAGAGATGACAGCTTACCAAGCCAGAACAAACGATGAGGCTCTGAGAAC	3177
Db	3121	ATCACGGAACGAGAGATGACAGCTTACCAAGCCAGAACAAACGATGAGGCTCTGAGAAC	3180
OY	3178	ACGTGCACATGCTGAGAGAACAGGTGATGATTTGAGAGCCCTTAACGATGAGCTGCTA	3237
Db	3181	ACGTGCACATGCTGAGAGAACAGGTGATGATTTGAGAGCCCTTAACGATGAGCTGCTA	3240
OY	3238	GAAATAAAGCGGCGAGTGGAGGCGCTGAGAGAGCGCTCTGGGTGATGAGAAATCCCAATTT	3297
Db	3241	GAAATAAAGCGGCGAGTGGAGGCGCTGAGAGAGCGCTCTGGGTGATGAGAAATCCCAATTT	3300
OY	3298	GAGTGTGGGTTGAGAGCTGACAGAGATGCTGACACCGAATAACAGACAGGCGAGAGA	3357
Db	3301	GAGTGTGGGTTGAGAGCTGACAGAGATGCTGACACCGAATAACAGACAGGCGAGAGA	3360
OY	3358	GCCGATCAAGCGGATCAACCGAGTCTCGCCAGGTGATGACTGGGACGTGAAGAGACACAAAG	3417
Db	3361	GCCGATCAAGCGGATCAACCGAGTCTCGCCAGGTGATGACTGGGACGTGAAGAGAGACACAAAG	3420
OY	3418	GCTGAGATTCCGCTCTGCACAGAGCTCTCAAAAGAGAGAACTGAGAGCCGAGAGCCTC	3477
Db	3421	GCTGAGATTCCGCTCTGCACAGAGCTCTCAAAAGAGAGAACTGAGAGCCGAGAGCCTC	3480
OY	3478	TCTGCACAGCTCAATGACCTGAGAGAGAGCATGCTATGCTTGAATGAATGCCCGAGAC	3537
Db	3481	TCTGCACAGCTCAATGACCTGAGAGAGAGCATGCTATGCTTGAATGAATGCCCGAGAC	3540
OY	3538	TTTACAGCAAGAGCTGAGAGCTGAACGAGAGCTCAAAACAGAGCTTCTGGAAGAGCAAGCC	3597
Db	3541	TTTACAGCAAGAGCTGAGAGCTGAACGAGAGCTCAAAACAGAGCTTCTGGAAGAGCAAGCC	3600
OY	3598	AAATTTACAGACGAGATGAGCTGACAGAAAATACATATTTCCGTCTGACCTCAAGGACTG	3657
Db	3601	AAATTTACAGACGAGATGAGCTGACAGAAAATACATATTTCCGTCTGACCTCAAGGACTG	3660
OY	3658	CAGAAGCTCTAGATCCGGCTGATCTTACTGAGACAGAAAAGAGTGACTTGGAGTATCAG	3717
Db	3661	CAGAAGCTCTAGATCCGGCTGATCTTACTGAGACAGAAAAGAGTGACTTGGAGTATCAG	3720
OY	3718	CTGGAACAACTTCAGGTTCTCTATTTCTCATGAAAAGGTGAATAATGAAAGGCACTATTTCT	3777
Db	3721	CTGGAACAACTTCAGGTTCTCTATTTCTCATGAAAAGGTGAATAATGAAAGGCACTATTTCT	3780
OY	3778	CAACAAACCAAACTCATTTGATTTTCTGACAGCCAAATGAGCCAACTGCTTAAAGAA	3837
Db	3781	CAACAAACCAAACTCATTTGATTTTCTGACAGCCAAATGAGCCAACTGCTTAAAGAA	3840
OY	3838	AAGGTTCCCTGCAGTACATGAGCTAGCTGAGCCCTGAGAGAGGAAAGCTCGCTG	3897

QY 6058 TCCCTCCGAGGCTGTTGAAGACAGACAGAGGCGGCTGCTGCGGAGCCGTGAGG 6117
DB 6061 TCCCTCCGAGGCTGTTGAAGACAGACAGAGGCGGCTGCTGCGGAGCCGTGAGG 6120
QY 6118 ACCCTCCGCTGCTCCAGGTGAACAGGTCTGGAGCACATC 6155
DB 6121 ACCCTCCGCTGCTCCAGGTGAACAGGTGAACAGCACATTC 6158
RESULT 8
ABS63436
ID ABS63436 standard; cDNA; 6189 BP.
XX
AC ABS63436;
XX
DT 15-NOV-2002 (first entry)
XX
DE RHO/RAC-interacting citron kinase-like human cDNA, designated NOV3b.
XX
DE Human; gene; ss; NOV; NOVX-associated disorder; cardiomyopathy;
KM atherosclerosis; diabetes; cell signalling; metabolic pathway;
KM cellular receptor; downstream effector; cancer; gene therapy;
KM hypertension; congenital heart defect; aortic stenosis; obesity;
KM infectious disease; anorexia; Alzheimer's disease; Parkinson's disease;
KM neurodegenerative disorder; haemophilia; dyslipidemia; vaccine;
KM neuropathic disease; scleroderma; fertility; immunogen;
KM idiopathic thrombocytopenic purpura; graft versus host disease;
KM Crohn's disease; multiple sclerosis; cirrhosis; autoimmune disease;
KM systemic lupus erythematosus; asthma; arthritis; psoriasis; allergy;
KM stroke; anxiety; Leech-Nyman syndrome; schizophrenia; cerebellar ataxia;
KM pain; alcoholism; transgenic.
XX
OS Homo sapiens.
XX
PN WO200226826-A2.
XX
PD 04-APR-2002.
XX
PF 27-SEP-2001; 2001WO-US042336.
XX
PR 27-SEP-2000; 2000US-0235631P.
XX
PR 27-SEP-2000; 2000US-0235633P.
XX
PR 27-SEP-2000; 2000US-0235808P.
XX
PR 27-SEP-2000; 2000US-0236064P.
XX
PR 27-SEP-2000; 2000US-0236065P.
XX
PR 27-SEP-2000; 2000US-0236066P.
XX
PR 28-SEP-2000; 2000US-0236135P.
XX
PR 03-OCT-2000; 2000US-0237434P.
XX
PR 05-OCT-2000; 2000US-0238321P.
XX
PR 06-OCT-2000; 2000US-0238396P.
XX
PR 06-OCT-2000; 2000US-0238398P.
XX
PR 16-MAR-2001; 2001US-0276667P.
XX
PR 31-MAY-2001; 2001US-0294823P.
XX
PR 12-JUL-2001; 2001US-0304868P.
XX
PR 26-SEP-2001; 2001US-00235631.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Gerlach VL, MacDougall JR, Smithson G, Millet I, Stone D;
PI Gunther E, Ellerman K, Groese WM, Alsbrook JP, Lepley DM;
PI Burgess CE, Padigaru M, Kekuda R, Spytek KA, Leach MD, Shinkets RA;
XX
DR WPI; 2002-489860/53.
XX
DR P-PSDB; ABG78363.
XX
PT Novel isolated NOVX polypeptides and polynucleotides homologous to
PT attractin, plexin, papin-like family of proteins, useful for treating
PT atherosclerosis, diabetes, cancer, Alzheimer's disease, hemophilia and
PT stroke.
XX
PS Claim 8; Page 43-44; 308pp; English.

CC The invention discloses the isolated human polypeptides, and
CC polynucleotides encoding them, that have been designated NOVX. The
CC polynucleotides, polynucleotides and antibodies are useful in treating or
CC preventing a NOVX-associated disorder which is cardiomyopathy,
CC atherosclerosis and diabetes in a human, where the disorder is related to
CC cell signal processing and metabolic pathway modulation. They can also be
CC used in determining the presence of, or predisposition to, a disease
CC associated with altered levels of the polypeptides and polynucleotides of
CC any one of the 13 sequences (NOV1-NOV8), for raising antibodies, for
CC identifying an agent that binds to, or that modulates the expression or
CC activity of the polypeptide, for identifying an agent which is cellular
CC receptor or downstream effector, for treating or preventing a NOVX-
CC associated disorder and as a pharmaceutical composition comprising the
CC polypeptide, polynucleotide or the antibody. The polypeptides and
CC polynucleotides are useful in diagnostic applications (e.g. as a marker
CC for cancerous cells or tissue types) where their amounts are assessed, or
CC for the manufacture of a medicament (e.g. gene therapy) for treating or
CC preventing disorders or syndromes such as hypertension, congenital heart
CC defects, aortic stenosis, obesity, infectious disease, anorexia, cancer,
CC Alzheimer's disease, Parkinson's disorder, neurodegenerative disorders,
CC haemophilia, dyslipidemias, haematopoietic diseases, scleroderma,
CC fertility, idiopathic thrombocytopenic purpura, graft versus host
CC diseases, Crohn's disease, multiple sclerosis, cirrhosis, autoimmune
CC disease, systemic lupus erythematosus, asthma, arthritis, psoriasis,
CC allergy, stroke, anxiety, Leech-Nyman syndrome, schizophrenia, cerebellar
CC ataxia, pain and alcoholism. They may also be used as immunogens to
CC produce antibodies specific for the invention, and as vaccines.
CC Transgenic cells containing a NOVX expressing construct are useful to
CC produce non-human transgenic animals for studying the function and/or
CC activity of the NOVX proteins and for identifying and/or evaluating
CC modulators of NOVX protein activity. Transgenic cells containing a NOVX
CC expressing construct are useful to produce non-human transgenic animals
CC for studying and/or evaluating modulators of NOVX proteins and for
CC identifying and/or evaluating modulators of NOVX protein activity. The
CC sequences presented in ABS63431-ABS63444 are the human NOV1-NOV8 cDNAs
XX
SQ Sequence 6189 BP; 1742 A; 1554 C; 1690 G; 1203 T; 0 U; 0 Other;
XX
Query Match 99.3%; Score 6121.8; DB 6; Length 6189;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 6145; Conservative 0; Mismatches 17; Indels 3; Gaps 1;
QY 1 ATGTGAAGTCAATATATGAGGCGGAAATCTTTGAGTCTGTGTAACCAATT 60
DB 1 ATGTGAAGTCAATATATGAGGCGGAAATCTTTGAGTCTGTGTAACCAATT 60
QY 61 GCCAGCCGGGCTCTCCAGGCTGAATCTGTTTCCAGGGGAAACCACTTTATGACTCAA 120
DB 61 GCCAGCCGGGCTCTCCAGGCTGAATCTGTTTCCAGGGGAAACCACTTTATGACTCAA 120
QY 121 CAGCAGATGCTCTCTTCCGAGAGGATTTAGATGCCCTTGTCTCTTTGAA 180
DB 121 CAGCAGATGCTCTCTTCCGAGAGGATTTAGATGCCCTTGTCTCTTTGAA 180
QY 181 GAATGAGTCAAGCTCTGTATGAAATTAACAGTGAACAATTGTCCGAAATAT 240
DB 181 GAATGAGTCAAGCTCTGTATGAAATTAACAGTGAACAATTGTCCGAAATAT 240
QY 241 TCCGACACCATTAAGTGAATTAACAGCTCCAGCTTCGGCAAGCACTTGAAGTACA 300
DB 241 TCCGACACCATTAAGTGAATTAACAGCTTCAGGCTTCGCAAGCACTTGAAGTACA 300
QY 301 AGCTTGTAAGTGTGTGCTACTTGTGTAAGTCAAGTGTGAAGAGAAAGCAACCGGG 360
DB 301 AGCTTGTAAGTGTGTGCTACTTGTGTAAGTCAAGTGTGAAGAGAAAGCAACCGGG 360
QY 361 GACATCTATGCTATGAAGTGAAGAAAGAGCTTTATTTGCCCCAGAGAGTTTCA 420
DB 361 GACATCTATGCTATGAAGTGAAGAAAGAGCTTTATTTGCCCCAGAGAGTTTCA 420
QY 421 TTTTGTGAGGAAGAGGGAACATTTATCTCGAAGCAAGCCCGTGGATCCCCCAATTA 480
DB 421 TTTTGTGAGGAAGAGGGAACATTTATCTCGAAGCAAGCCCGTGGATCCCCCAATTA 480

```
QY 481 CAGTATGCTTTTCAGAGCAAAATCACTTTATCTGCTCATGGAATATCAGCTTGAGGG 540
DB 481 CAGTATGCTTTTCAGAGCAAAATCACTTTATCTGCTCATGGAATATCAGCTTGAGGG 540
QY 541 GACTTGTCTGCTTTTGAATATGAGGACCAAGTTAGATGAATCTGATCAGTTT 600
DB 541 GACTTGTCTGCTTTTGAATATGAGGACCAAGTTAGATGAATCTGATCAGTTT 600
QY 601 TACTTACCTGAGCTGATTTTGGCTGTTTCAAGCGTTCACTGATGGATACGTCATCGA 660
DB 601 TACTTACCTGAGCTGATTTTGGCTGTTTCAAGCGTTCACTGATGGATACGTCATCGA 660
QY 661 GAGATCAAGCTTGAGCAATCTGTTGACCGGACAGGACATCAAGCTGTGTGATTTT 720
DB 661 GAGATCAAGCTTGAGCAATCTGTTGACCGGACAGGACATCAAGCTGTGTGATTTT 720
QY 721 GGAATCTCCGCGAAATGAATTTCAACAAAGATGATGCAATCCGGAATTGGAGCC 780
DB 721 GGAATCTCCGCGAAATGAATTTCAACAAAGATGATGCAATCCGGAATTGGAGCC 780
QY 781 CCAAGATTACATGCTCTCTGAAAGTCTGATCTGTGATGAACGGGATGGAAGGCACTTAC 840
DB 781 CCAAGATTACATGCTCTCTGAAAGTCTGATCTGTGATGAACGGGATGGAAGGCACTTAC 840
QY 841 GGGCTGGAATCTGATCTGTGCTGATGAGTGGGCTGATTTGCTTATGATGATTTATGGAGA 900
DB 841 GGGCTGGAATCTGATCTGTGCTGATGAGTGGGCTGATTTGCTTATGATGATTTATGGAGA 900
QY 901 TCCCTCTGCGAGAGGACCTCTGCGAGAACCTTCAATTAACATTATGAATTTCCAGCGG 960
DB 901 TCCCTCTGCGAGAGGACCTCTGCGAGAACCTTCAATTAACATTATGAATTTCCAGCGG 960
QY 961 TTTTGAATTTTCCAGATGACCCCAAGATGAGCACTTCTTGAATCTGATTTCAAGC 1020
DB 961 TTTTGAATTTTCCAGATGACCCCAAGATGAGCACTTCTTGAATCTGATTTCAAGC 1020
QY 1021 TTGTTGTGGGCGCAAGAAAGAGATGAAAGTTGAAAGTCTTTGCTGCACTTTCTTC 1080
DB 1021 TTGTTGTGGGCGCAAGAAAGAGATGAAAGTTGAAAGTCTTTGCTGCACTTTCTTC 1080
QY 1081 TCTTAAATTTGATGAGAAACATTTGTAATCTCTCTCCCTCCCTGCTCCACCTTCAAG 1140
DB 1081 TCTTAAATTTGATGAGAAACATTTGTAATCTCTCTCCCTCCCTGCTCCACCTTCAAG 1140
QY 1141 TCTGAGATGACACTTCAATTTTGAATGAACAGAGAAAGATTTGATGATTTTCTCTCT 1200
DB 1141 TCTGAGATGACACTTCAATTTTGAATGAACAGAGAAAGATTTGATGATTTTCTCTCT 1200
QY 1201 CCGTGCACGTGAGCCCTTCAAGCTTCTGCGGTGAAGAACTGCGCTTTGTGGGGTTTTCG 1260
DB 1201 CCGTGCACGTGAGCCCTTCAAGCTTCTGCGGTGAAGAACTGCGCTTTGTGGGGTTTTCG 1260
QY 1261 TACAGCAAGGCACTGAGGATTTCTTGTAGATCTGATCTGTTGTGTGCGGTCTGAGCTCC 1320
DB 1261 TACAGCAAGGCACTGAGGATTTCTTGTAGATCTGATCTGTTGTGTGCGGTCTGAGCTCC 1320
QY 1321 CCTGCGAAGACTAGCTCAATGAAAGAACTTCTCATCAAAAGCAAAAGACTTACAAGAC 1380
DB 1321 CCTGCGAAGACTAGCTCAATGAAAGAACTTCTCATCAAAAGCAAAAGACTTACAAGAC 1380
QY 1381 TCTCAGAGCAAGTGTCAAGAATGAGCAAGAAATGACCCGGTTACATCGAGAGTGTCA 1440
DB 1381 TCTCAGAGCAAGTGTCAAGAATGAGCAAGAAATGACCCGGTTACATCGAGAGTGTCA 1440
QY 1441 GAGGTGAGGCTGTGCTTATGTCAGAAAGAGGTGAGCTGAAGCCCTCTGAGACTCAGANA 1500
DB 1441 GAGGTGAGGCTGTGCTTATGTCAGAAAGAGGTGAGCTGAAGCCCTCTGAGACTCAGANA 1500
QY 1501 TCCCTCTGAGCAGGACCTTGTCTACTTACATCAAGAAATGCAAGTGTGTTAAAGCGAAGT 1560
DB 1501 TCCCTCTGAGCAGGACCTTGTCTACTTACATCAAGAAATGCAAGTGTGTTAAAGCGAAGT 1560
QY 1561 TTGAGCAAGACCGAGATGAGGTGTCCAGAGAGATGACAAAGCACTGACAGCTTCTCAT 1620
DB 1561 TTGAGCAAGACCGAGATGAGGTGTCCAGAGAGATGACAAAGCACTGACAGCTTCTCAT 1620
QY 1621 GATATCAGAGACGAGCCGAGAGCTCCAGAAATCAAAAGAGAGAGTACAGGCTTCA 1680
DB 1621 GATATCAGAGACGAGCCGAGAGCTCCAGAAATCAAAAGAGAGAGTACAGGCTTCA 1680
QY 1681 GTGAGAAATATGAGTTGATGATGAATCAAGTTGGAAGAGATCTTGTCTCAGCAAGAGA 1740
DB 1681 GTGAGAAATATGAGTTGATGATGAATCAAGTTGGAAGAGATCTTGTCTCAGCAAGAGA 1740
QY 1741 CGAGTGTATCTTACGAATCTGAGCTGAGAGAGTCTCGGCTTGTCTGTAAGAAATTCAG 1800
DB 1741 CGAGTGTATCTTACGAATCTGAGCTGAGAGAGTCTCGGCTTGTCTGTAAGAAATTCAG 1800
QY 1801 CGGAAAGCGACAGAAATGTCAGCATTAATCTGTAAGCTGTAAGATCAAGGAAAGCTTGA 1860
DB 1801 CGGAAAGCGACAGAAATGTCAGCATTAATCTGTAAGCTGTAAGATCAAGGAAAGCTTGA 1860
QY 1861 GTGAGAAATATGAGTTGATGATGAATCAAGTTGGAAGAGATCTTGTCTCAGCAAGAGA 1920
DB 1861 GTGAGAAATATGAGTTGATGATGAATCAAGTTGGAAGAGATCTTGTCTCAGCAAGAGA 1920
QY 1921 CTCCAAGAGAACTGAGAAAGGCTGTAAAGCCAGACCGAGGCTTGTGAG 1980
DB 1921 CTCCAAGAGAACTGAGAAAGGCTGTAAAGCCAGACCGAGGCTTGTGAG 1980
QY 1981 AATATCCGCGAGCAAAAGAGCGAGCGAGAGGAGCTGAGAAAGCTTGCAGAACCCGAG 2040
DB 1981 AATATCCGCGAGCAAAAGAGCGAGCGAGAGGAGCTGAGAAAGCTTGCAGAACCCGAG 2040
QY 2041 GATTTCTTGAAGGCTATCAGAAAGAGAGCTGTGAGAGTGAAGAGCGCCGCTTCTCTG 2100
DB 2041 GATTTCTTGAAGGCTATCAGAAAGAGAGCTGTGAGAGTGAAGAGCGCCGCTTCTCTG 2100
QY 2101 GAGAACAGGTAAGAGATGATGAGACCATGAGAGCTGAGAGAAACAGACTGAAGATGAC 2160
DB 2101 GAGAACAGGTAAGAGATGATGAGACCATGAGAGCTGAGAGAAACAGACTGAAGATGAC 2160
QY 2161 ATCCAGCAAAATCCCAAGATCCAGAGATGAGCTGATTAATTTCTGAGCTTGAAGAG 2220
DB 2161 ATCCAGCAAAATCCCAAGATCCAGAGATGAGCTGATTAATTTCTGAGCTTGAAGAG 2220
QY 2221 AAAATCGGAGGCGCAAGTCTCAGCCAGCACTTGAAGTGCACCTGAAACAGAAAGAG 2280
DB 2221 AAAATCGGAGGCGCAAGTCTCAGCCAGCACTTGAAGTGCACCTGAAACAGAAAGAG 2280
QY 2281 CAGCACTATGAGAAAGATTAAGTGTGGAACAATCAGATTAAGAAAGACCTGCTGAC 2340
DB 2281 CAGCACTATGAGAAAGATTAAGTGTGGAACAATCAGATTAAGAAAGACCTGCTGAC 2340
QY 2341 AAGAGAGCACTGAGAAACATGATGAGAGACACAGAGAGAGAGGCCATGAGAAAGGCAAA 2400
DB 2341 AAGAGAGCACTGAGAAACATGATGAGAGACACAGAGAGAGAGGCCATGAGAAAGGCAAA 2400
QY 2401 ATTCTCAGCGAAACAGAGGCGATGATCAATGCTATGAGATTTCAAGATCAGATCCCTGAA 2460
DB 2401 ATTCTCAGCGAAACAGAGGCGATGATCAATGCTATGAGATTTCAAGATCAGATCCCTGAA 2460
QY 2461 CAGAGATTTGTGAACTGTCTGAGCCAAATAACTTGCAGCAATAGCAGTCTTTTTCAC 2520
DB 2461 CAGAGATTTGTGAACTGTCTGAGCCAAATAACTTGCAGCAATAGCAGTCTTTTTCAC 2520
QY 2521 CAAAAGAAATATGAGGCGCCAGAGAGAGATTTTCTGAACTCAGGCAACAGAAATTTTAC 2580
DB 2521 CAAAAGAAATATGAGGCGCCAGAGAGAGATTTTCTGAACTCAGGCAACAGAAATTTTAC 2580
QY 2581 CTGAGAGACAGGCTGTGAGAGGCGCCAGAACCGAGAACTGAGAGAGAGAGCTGAGAG 2640
DB 2581 CTGAGAGACAGGCTGTGAGAGGCGCCAGAACCGAGAACTGAGAGAGAGAGCTGAGAG 2640
QY 2641 AAGATCAGCCACCAAGACCAAGTGAAGATCGGCTGTGAGACTGAGAGCAAGATTTG 2700
DB 2641 AAGATCAGCCACCAAGACCAAGTGAAGATCGGCTGTGAGACTGAGAGCAAGATTTG 2700
```

Db 2638 AAGATCCACCAAGACCAAGTACAGAAATCGCTGCACTGAGAACAGGATTG 2697
Qy 2701 CCGGAGCTCAGTCTAGAGCAGAGAGCAGAACTGAGCTCAAGCCAGCTCAGAG 2760
Db 2698 CCGGAGTGAAGTCTAGAGCAGAGAGCAGAACTGAGCTCAAGCCAGCTCAGAG 2757
Qy 2761 CTACAGCTCTCCCTGAGAGCGGAGTCAAGTTGAACGCTTCAGAGCTGACGGCG 2820
Db 2758 CTACAGCTCTCCCTGAGAGCGGAGTCAAGTTGAACGCTTCAGAGCTGACGGCG 2817
Qy 2821 GCCCTGAGAGCCAGCTTCCGAGAGCCAGAGACAGAGCTGAGAGACCAAGAGCT 2880
Db 2818 GCCCTGAGAGCCAGCTTCCGAGAGCCAGAGACAGAGCTGAGAGACCAAGAGCT 2877
Qy 2881 GAAGAGAGATCCAGGCACTCAGGCACTAGAGATGAATCCAGCCAAATTTGATGCT 2940
Db 2878 GAAGAGAGATCCAGGCACTCAGGCACTAGAGATGAATCCAGCCAAATTTGATGCT 2937
Qy 2941 CTTGTAACAGCTGTACTGTATTCAGACCTGAGAGAGAGCTAAACAGCTGACCGAG 3000
Db 2938 CTTGTAACAGCTGTACTGTATTCAGACCTGAGAGAGAGCTAAACAGCTGACCGAG 2997
Qy 3001 GACAACTGTAATCTCAACAAACAACTTCTACTTGTCCAAACAATCGATGAGGCTTCT 3060
Db 2998 GACAACTGTAATCTCAACAAACAACTTCTACTTGTCCAAACAATCGATGAGGCTTCT 3057
Qy 3061 GGGGCCAAGACGAGATTGTACAACTGAGAGTGAAGTGAACATCTCCGCGGAGATC 3120
Db 3058 GGGGCCAAGACGAGATTGTACAACTGAGAGTGAAGTGAACATCTCCGCGGAGATC 3117
Qy 3121 ACCGAACGAGAGATGACGCTTACAGCAGAGAGCAACGATGAGGCTGAGAGACGAG 3180
Db 3118 ACCGAACGAGAGATGACGCTTACAGCAGAGAGCAACGATGAGGCTGAGAGACGAG 3177
Qy 3181 TGCACTGCTGAGAGAACAGTCAATGATTTGAGGCTTAAACATGAGCTCTAGAA 3240
Db 3178 TGCACTGCTGAGAGAACAGTCAATGATTTGAGGCTTAAACATGAGCTCTAGAA 3237
Qy 3241 AAAGAGCGGAGTGGAGGCTGAGAGAGCTCTGGGTGATGAGAAATCCCACTTGAAG 3300
Db 3238 AAAGAGCGGAGTGGAGGCTGAGAGAGCTCTGGGTGATGAGAAATCCCACTTGAAG 3297
Qy 3301 TGTGCGTTGAGAGCTGAGAGAGCTGAGACACGAGAAACAGAGCGGAGAGAGCC 3360
Db 3298 TGTGCGTTGAGAGCTGAGAGAGCTGAGACACGAGAAACAGAGCGGAGAGAGCC 3357
Qy 3361 GATCAGCGGATCACGAGTCTCCAGGTGCTGAGCTGAGCTGAGAGAGCAAGGCT 3420
Db 3358 GATCAGCGGATCACGAGTCTCCAGGTGCTGAGCTGAGCTGAGAGAGCAAGGCT 3417
Qy 3421 GAGATTTCCCTCTGAGAGAGAGCTCTCAAAAGAGAGAGCTGAGAGAGCTCTCT 3480
Db 3418 GAGATTTCCCTCTGAGAGAGAGCTCTCAAAAGAGAGAGCTGAGAGAGCTCTCT 3477
Qy 3481 GACAACTCATGACTGAGAGAGAGAGATGCTGAGTGAATGAATGCCAAGCTTA 3540
Db 3478 GACAACTCATGACTGAGAGAGAGAGATGCTGAGTGAATGAATGCCAAGCTTA 3537
Qy 3541 CAGCAGAGCTGAGAGCTGAACGAGACTCAACAGAGCTTCTGAGAGAGCAAGCCAA 3600
Db 3538 CAGCAGAGCTGAGAGCTGAACGAGACTCAACAGAGCTTCTGAGAGAGCAAGCCAA 3597
Qy 3601 TTACAGAGAGAGTGAACCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3660
Db 3598 TTACAGAGAGAGTGAACCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3657
Qy 3661 GAAGCTCTAGATCGGAGCTGATCTACTGAAGACAGAGAGAGAGAGAGAGAGAGAGAG 3720
Db 3658 GAAGCTCTAGATCGGAGCTGATCTACTGAAGACAGAGAGAGAGAGAGAGAGAGAGAG 3717
Qy 3721 GAAACATTCAGGTTCTCTATTTCTATGAAAGTGAAGAGAGAGAGAGAGAGAGAGAG 3780

Db 3718 GAAACATTCAGGTTCTCTATTTCTATGAAAGTGAAGAGAGAGAGAGAGAGAGAG 3777
Qy 3781 CAACCAACTCTATGATTTTCTGCAAGCCAAATGAGCAACCTGCTTAAAAAGAAAAAG 3840
Db 3778 CAACCAACTCTATGATTTTCTGCAAGCCAAATGAGCAACCTGCTTAAAAAGAAAAAG 3837
Qy 3841 GTTCTCTGAGATCAATGAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3900
Db 3838 GTTCTCTGAGATCAATGAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3897
Qy 3901 GAGCTAG 3960
Db 3898 GAGCTAG 3957
Qy 3961 GCCCAGCGCAAG 4020
Db 3958 GCCCAGCGCAAG 4017
Qy 4021 ATGCGCATGTCGCTCATGTCGCTGCGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4080
Db 4018 ATGCGCATGTCGCTCATGTCGCTGCGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4077
Qy 4081 GCCCGGCAATCCAGCGGAG 4140
Db 4078 GCCCGGCAATCCAGCGGAG 4137
Qy 4141 AAGGAACGATGACACCAATATTTCTCAAGATTCACAGTGAAGTGAAGTGAAGTGAAG 4200
Db 4138 AAGGAACGATGACACCAATATTTCTCAAGATTCACAGTGAAGTGAAGTGAAGTGAAG 4197
Qy 4201 ACAAAGT 4260
Db 4198 ACAAAGT 4257
Qy 4261 GAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4320
Db 4258 GAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4317
Qy 4321 CTTGCTGAATATGCAACACTTCAAGAGAGCTTCTGCGGTGACAAATGAAATGAACTCCCA 4380
Db 4318 CTTGCTGAATATGCAACACTTCAAGAGAGCTTCTGCGGTGACAAATGAAATGAACTCCCA 4377
Qy 4381 GGTCTCAGACCAAG 4440
Db 4378 GGTCTCAGACCAAG 4437
Qy 4441 AGGAATTAACAAAG 4500
Db 4438 AGGAATTAACAAAG 4497
Qy 4501 AAGTCTCAATTAATGAATGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4560
Db 4498 AAGTCTCAATTAATGAATGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4557
Qy 4561 GAGCTGTGCTTCCGAG 4620
Db 4558 GAGCTGTGCTTCCGAG 4617
Qy 4621 GCAAAATCAGCAG 4680
Db 4618 GCAAAATCAGCAG 4677
Qy 4681 ACCGTGAGGCGCGGAG 4740
Db 4678 ACCGTGAGGCGCGGAG 4737
Qy 4741 TGAGTCAACGCTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4800
Db 4738 TGAGTCAACGCTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4797
Qy 4801 GCTGATGCTAACTGCTTGAAG 4860
Db 4798 GCTGATGCTAACTGCTTGAAG 4857

QY	4861	ATGAAC	TGACACCGCTCCCTTCA	GTGACACGAGTGTGTGTGTGGGCA	CCGAGAAAGGCTC	4920
Db	4858	ATGAAC	TGACACCGCTCCCTTCA	GTGACACGAGTGTGTGTGTGGGCA	CCGAGAAAGGCTC	4917
QY	4921	TACGCC	CTGAAGTCTTGTAAAAAC	TCCCTAACCCATGTCCAGAAATTGGAGCA	GTCTTC	4980
Db	4918	TACGCC	CTGAAGTCTTGTAAAAAC	TCCCTAACCCATGTCCAGAAATTGGAGCA	GTCTTC	4977
QY	4961	CAAA	TTTATATATATCAAGSAC	CTTGAGAACTTCAATGATAGCAGAAAGACGGGCA	5040	
Db	4978	CAAA	TTTATATATATCAAGSAC	CTTGAGAACTTCAATGATAGCAGAAAGACGGGCA	5037	
QY	5041	CTGTGT	CTTTGTGAGAGTGAAGAA	AGTGAACGTCTCCGCGCCAGTCCCATCTGCGC	5100	
Db	5038	CTGTGT	CTTTGTGAGAGTGAAGAA	AGTGAACGTCTCCGCGCGAGTCCCATCTGCGC	5097	
QY	5101	CAGCCG	CAATCTCAACCCAACTTTT	TAAGTGTCAAGGGCTGCACATTTTGGGGCA	5160	
Db	5098	CAGCCG	CAATCTCAACCCAACTTTT	TAAGTGTCAAGGGCTGCACATTTTGGGGCA	5157	
QY	5161	GGCA	GATTTGAGAACGGGCTCTG	CACTGTGCAAGCCAGACCAAGTGTCTATTC	5220	
Db	5158	GGCA	GATTTGAGAACGGGCTCTG	CACTGTGCAAGCCAGACCAAGTGTCTATTC	5217	
QY	5221	CGCTA	CAACGAACCACTCAGCA	AATACTGCATCCGAAAAGATAGACCTCAGAGCC	5280	
Db	5218	CGCTA	CAACGAACCACTCAGCA	AATACTGCATCCGAAAAGATAGACCTCAGAGCC	5277	
QY	5281	TGCAGCT	GTATCCACTTCAACCA	ATTACGTATTCCTCATTTGGAACCAATTTTACGAA	5340	
Db	5278	TGCAGCT	GTATCCACTTCAACCA	ATTACGTATTCCTCATTTGGAACCAATTTTACGAA	5337	
QY	5341	ATCGA	CATGAAGACGTACACGCT	CGAGGAATTCCGTGAATAGAAATGACATTTCTTGGA	5400	
Db	5338	ATCGA	CATGAAGACGTACACGCT	CGAGGAATTCCGTGAATAGAAATGACATTTCTTGGA	5397	
QY	5401	CTGTGT	GTGTGTGCCGCTCTT	CCAACAGCTTCCCTGTCTCAATCTGTGAGAGTAACAGC	5460	
Db	5398	CTGTGT	GTGTGTGCCGCTCTT	CCAACAGCTTCCCTGTCTCAATCTGTGAGAGTAACAGC	5457	
QY	5461	GCAGGG	CAGCGAGAGTACTTGT	GTGTTCACGAATTTGGAAGTTCGTGTGTCT	5520	
Db	5458	GCAGGG	CAGCGAGAGTACTTGT	GTGTTCACGAATTTGGAAGTTCGTGTGTCT	5517	
QY	5521	TACG	GAAGACGTAGCCGACAG	AGATCTCAAGTGAAGTGTGCTTACCTTTGGCTTTGCC	5580	
Db	5518	TACG	GAAGACGTAGCCGACAG	AGATCTCAAGTGAAGTGTGCTTACCTTTGGCTTTGCC	5577	
QY	5581	TACAG	AAGAACCTTATCTGT	TGTGTGACCACTTCACTACATCGAAGTAAATTGAGATCAG	5640	
Db	5578	TACAG	AAGAACCTTATCTGT	TGTGTGACCACTTCACTACATCGAAGTAAATTGAGATCAG	5637	
QY	5641	GCAC	GTCTCTAGCAGGAGAC	CCCTGTGCCGAGCGTACTTGACATCCGAAACCCGCGCTAC	5700	
Db	5638	GCAC	GTCTCTAGCAGGAGAC	CCCTGTGCCGAGCGTACTTGACATCCGAAACCCGCGCTAC	5697	
QY	5701	CTGG	GGCCCTGCACTTCTCT	CAAGAGCAATTAATTGCGGTCTCTATATCCAGATTAATTA	5760	
Db	5698	CTGG	GGCCCCCTGCACTTCTCT	CAAGAGCAATTAATTGCGGTCTCTATATCCAGATTAATTA	5757	
QY	5761	AGGT	CTATTTTCTTCAAGGAA	AACCTGTGTGAAGATCCGGCACTGAACACCAACCGGGG	5820	
Db	5758	AGGT	CTATTTTCTTCAAGGAA	AACCTGTGTGAAGATCCGGCACTGAACACCAACCGGGG	5817	
QY	5821	CCGT	CCACTTCCGAGACAG	CCCCAACAGGAGGCCACCACTACGATACAGACATC	5880	
Db	5818	CCGT	CCACTTCCGAGACAG	CCCCAACAGGAGGCCACCACTACGATACAGACATC	5877	
QY	5881	ACCA	AGCGCGTGTCTTCAAG	CCCAAGCCGCGCGCGAGAAAGGCCCAACCCGCGAGAGCCA	5940	
Db	5878	ACCA	AGCGCGCGTGTCTTCAAG	CCCAAGCCGCGCGCGCGAGAAAGGCCCAACCCGCGAGAGCCA	5937	

QY	5941	AGCACCACCCCAACCGCTACCGGAGAGGGCGGACCCGAGCTGCCAGGAGACAACTCTCTGAC	6000
Db	5938	AGCACAACCCCAACCGCTACCGGAGAGGGCGGACCCGAGCTGCCAGGAGACAACTCTCTGAC	5997
QY	6001	CGCCCCCTGGAGCGGAGAACTCCCCCGGACCGGATACCTCAGCAGCCGAGAGAGCGGTC	6066
Db	5998	CGCCCCCTGGAGCGGAGAACTCCCCCGGACCGGATGCTCAGCAGCCGAGAGAGCGGTC	6057
QY	6061	CCCGGAGGCTGTTTGAAGACAGCAGCAGGGGCGGCTGCGGAGCCGTGAGAC	6120
Db	6058	CCCGGAGGCTGTTTGAAGACAGCAGCAGGGGCGGCTGCGGAGCCGTGAGAC	6117
QY	6121	CCGCTGTCCAGGTGACACAGGCTGTGGACCAAGTCTTCAATATA	6165
Db	6118	CCGCTGTCCAGGTGACACAGGCTGTGGACCAAGTCTTCAATATA	6162
RESULT 9			
ID	ADA05641	standard; cDNA; 6189 BP.	
XX	ADA05641;		
AC	ADA05641;		
XX	06-NOV-2003	(first entry)	
DE	Human NOVA1 encoding cDNA SEQ ID NO:1.		
XX	human; NOVA1; antidiabetic; anorectic; antibacterial; virucide;		
KW	immunomodulator; cytoprotatic; neurotropic; neuroprotective;		
KW	antiparkinsonian; antilipemic; gene therapy; human disease;		
KW	metabolic disorder; diabetes; obesity; infection; cachexia; cancer;		
KW	neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;		
XX	immune disorder; haematopoietic disorder; dyslipidaemia; gene; ss.		
XX	Homo sapiens.		
XX	Key	Location/Qualifiers	
FT	CDS	1..6162	
FT		/*tag= a	
FT		/product= "NOVA1"	
XX	WO2003029424-A2.		
PD	10-APR-2003.		
PF	02-OCT-2002; 2002WO-US031373.		
XX	02-OCT-2001; 2001US-0326483P.		
PR	05-OCT-2001; 2001US-0327435P.		
PR	05-OCT-2001; 2001US-0327449P.		
PR	09-OCT-2001; 2001US-0327917P.		
PR	09-OCT-2001; 2001US-0328029P.		
PR	09-OCT-2001; 2001US-0328044P.		
PR	09-OCT-2001; 2001US-0328056P.		
PR	12-OCT-2001; 2001US-0328849P.		
PR	15-OCT-2001; 2001US-0329414P.		
PR	17-OCT-2001; 2001US-0330142P.		
PR	18-OCT-2001; 2001US-0330309P.		
PR	22-OCT-2001; 2001US-0341058P.		
PR	24-OCT-2001; 2001US-0343626P.		
PR	24-OCT-2001; 2001US-0343629P.		
PR	29-OCT-2001; 2001US-0349575P.		
PR	01-NOV-2001; 2001US-0346357P.		
PR	17-APR-2002; 2002US-0373260P.		
PR	19-APR-2002; 2002US-0373815P.		
PR	19-APR-2002; 2002US-0373817P.		
PR	19-APR-2002; 2002US-0373826P.		
PR	19-APR-2002; 2002US-0373884P.		
PR	22-APR-2002; 2002US-0374977P.		
PR	16-MAY-2002; 2002US-0381037P.		
PR	16-MAY-2002; 2002US-0381038P.		
PR	16-MAY-2002; 2002US-0381042P.		
PR	17-MAY-2002; 2002US-0381542P.		

PR 28-MAY-2002; 2002US-0383656P.
PR 29-MAY-2002; 2002US-0383831P.
PR 25-JUN-2002; 2002US-0391335P.
PR 01-OCT-2002; 2002US-00262511.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Smithson G, Millet I, Peyman JA, Kekuda R, Ju J, Li L, Guo X;
PI Patlurajan M, Spytek KA, Edinger SR, Ellerman K, Malayakar UM;
PI Ort T, Gorman L, Zehusen BD, Anderson DW, Zhong M, Catterton E;
PI Ji W, Miller CE, Rastelli L, Stone DJ, Pena CE, Shenoy SG;
PI Shinkete RA, Rothenberg ME, Leach MD, Agee ML, Bergins C, Dipippo VA;
PI Eisen AJ, Gangoli EA, Rieger DK, Spaderna SK;
XX
DR WPI; 2003-381626/36.
XX P-SDB; ADA05642.
XX
PT New NOX polypeptides and nucleic acids, useful for diagnosing,
PT preventing or treating NOX-associated disorders, e.g. diabetes, obesity,
PT cancer or dyslipidemia, and in chromosome mapping, tissue typing or
PT pharmacogenomics.
XX
PS Claim 20; Page 98-99; 586pp; English.
XX
XX The present invention describes NOX proteins, where X can be 1 to 55
CC (e.g. NOY1). Also described: (1) a composition comprising a polypeptide
CC described above and a carrier; (2) a kit comprising, in one or more
CC containers, the composition described above; (3) an isolated nucleic acid
CC molecule which encodes a NOX protein of the invention; (4) a vector
CC comprising the nucleic acid molecule described above; (5) a cell
CC comprising the above vector; (6) an antibody that immunospecifically
CC binds to the polypeptide described above; (7) methods for determining the
CC presence or amount of the above polypeptide or nucleic acid molecule in a
CC sample; (8) methods for determining the presence of or predisposition to
CC a disease associated with altered levels of expression of the above
CC polypeptide or nucleic acid molecule in a first mammalian subject; (9) a
CC method of identifying an agent that binds to the polypeptide described
CC above; (10) a method for identifying a potential therapeutic agent for
CC use in treating a pathology that is related to an aberrant expression or
CC aberrant physiological interactions of the polypeptide; (11) a method of
CC screening for a modulator of activity or of latency or predisposition to
CC a pathology associated with the polypeptide; (12) a method for modulating
CC the activity of the polypeptide described above; (13) methods of treating
CC or preventing a pathology associated with the above polypeptide in a
CC mammal; and (14) a method for producing the above polypeptide. NOX
CC sequences have antidiabetic, anorectic, antibacterial, virucide,
CC immunomodulator, cytototoxic, neurotropic, neuroprotective, antiparkinsonian
CC and antilipemic activities, and can be used in gene therapy. The
CC polypeptide is useful in manufacturing a medicament for treating a
CC syndrome associated with a human disease. The polypeptide or the nucleic
CC acid molecule may be used to diagnose, treat or prevent metabolic
CC disorders such as diabetes or obesity, infections, cachexia, cancer,
CC neurodegenerative disorders such as Alzheimer's disease or Parkinson's
CC disease, immune disorders, haematopoietic disorders and various
CC dyslipidemias. The nucleic acids can also be used as hybridisation
CC probes, in chromosome mapping, tissue typing, preventive medicine and
CC pharmacogenomics. The present sequence encodes a human NOX protein from
CC the present invention.
XX
SQ Sequence 6189 BP; 1742 A; 1554 C; 1690 G; 1203 T; 0 U; 0 Other;
Query Match 99.3%; Score 6121.8; DB 8; Length 6189;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 6145; Conservative 0; Mismatches 17; Indels 3; Gaps 1;
QY 1 ATGTGAAGTTCAAATATGAGCGCGGAATCTTGTGATGCTGTGCTGTAACCAATT 60
DB 1 ATGTGAAGTTCAAATATGAGCGCGGAATCTTGTGATGCTGTGCTGTAACCAATT 60
QY 61 GCGAGCGGCGCTCCAGGCTGAATCTGTTCTTCCAGGGGAAACACCCCTTATATCACTCAA 120
DB 61 GCGAGCGGCGCTCCAGGCTGAATCTGTTCTTCCAGGGGAAACACCCCTTATATCACTCAA 120

QY 121 CAGCAGATGCTCTCTCTTTCCCGAAGAGGATATTAATGCCCTTGTCTTGTGAA 180
DB 121 CAGCAGATGCTCTCTCTTTCCCGAAGAGGATATTAATGCCCTTGTCTTGTGAA 180
QY 181 GAATGAGTGAAGCTGCTGTGATGAAGATTAAGACGTGAGCACTTGTCCGGAAGTAT 240
DB 181 GAATGAGTGAAGCTGCTGTGATGAAGATTAAGACGTGAGCACTTGTCCGGAAGTAT 240
QY 241 TCCGACACCATAGCTGATGATGAGAGAGCTCCAGCTTCGCGCAAGACCTTGAAGTACA 300
DB 241 TCCGACACCATAGCTGATGATGAGAGAGCTCCAGCTTCGCGCAAGACCTTGAAGTACA 300
QY 301 AGCTCTGATGATGCTGTGCTCACTTCTGTAAGTGCAGGTGTGAAGAGCAACCGGG 360
DB 301 AGCTCTGATGATGCTGTGCTCACTTCTGTAAGTGCAGGTGTGAAGAGCAACCGGG 360
QY 361 GACATCATGCTATGAAAGATGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 420
DB 361 GACATCATGCTATGAAAGATGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 420
QY 421 TTTTGTGAG 480
DB 421 TTTTGTGAG 480
QY 481 CAGTATGCTTTTCAAGACAAAATACCTTATCTGTGATGAAATACGCTGAGAGG 540
DB 481 CAGTATGCTTTTCAAGACAAAATACCTTATCTGTGATGAAATACGCTGAGAGG 540
QY 541 GACTTGCTGCACTTTTGAATATGATATGAGACAGATTAATGAATCACTGATCAGTT 600
DB 541 GACTTGCTGCACTTTTGAATATGATATGAGACAGATTAATGAATCACTGATCAGTT 600
QY 601 TACCTAGCTAGCTGATTTTGGCTGTTCAAGAGCTTATCTGATGAGATACGTGATGCA 660
DB 601 TACCTAGCTAGCTGATTTTGGCTGTTCAAGAGCTTATCTGATGAGATACGTGATGCA 660
QY 661 GACATCAAGCTGAGAAATCTGCTGTTGACCGACAGACACATCAAGCTGTGATGATTT 720
DB 661 GACATCAAGCTGAGAAATCTGCTGTTGACCGACAGACACATCAAGCTGTGATGATTT 720
QY 721 GATCTGCGCGGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATG 780
DB 721 GATCTGCGCGGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATG 780
QY 781 CCGATTAATGAGCTCTGTAAGTCTGATGATGAGAGAGAGAGAGAGAGAGAGAGAG 840
DB 781 CCGATTAATGAGCTCTGTAAGTCTGATGATGAGAGAGAGAGAGAGAGAGAGAGAG 840
QY 841 GGCCTGACCTGATCTGAGTGTGATGAGTGTGATGAGTGTGATGAGTGTGATGAGTGT 900
DB 841 GGCCTGACCTGATCTGAGTGTGATGAGTGTGATGAGTGTGATGAGTGTGATGAGTGT 900
QY 901 TCCCTGCTGCAAG 960
DB 901 TCCCTGCTGCAAG 960
QY 961 TCCCTGCTGCAAG 997
DB 961 TCCCTGCTGCAAG 997
QY 997 TCCCTGCTGCAAG 1020
DB 997 TCCCTGCTGCAAG 1020
QY 1021 TTTTGTGAG 1080
DB 1021 TTTTGTGAG 1080
QY 1081 TTTTGTGAG 1140
DB 1081 TTTTGTGAG 1140
QY 1141 TTTTGTGAG 1200
DB 1141 TTTTGTGAG 1200
QY 1201 CCGTGCAGCTGAGCCCTCAGGCTTCTCGGAGTGAAGAGAGAGAGAGAGAGAGAGAG 1260
DB 1201 CCGTGCAGCTGAGCCCTCAGGCTTCTCGGAGTGAAGAGAGAGAGAGAGAGAGAGAG 1260

1198 CCGTGCAGCTGAGCCCTTCAAGGCTTCTCGGGTGAAGAACTGCCGTTTGTGGGGTTTTTCG 1257
QY 1261 TACAGCAAGGCATCGGGGATTTCTTGTAGATCTGAGTCTGTGTGTGTCGGGTCTGAGCTCC 1320
Db 1258 TACAGCAAGGCATCGGGGATTTCTTGTAGATCTGAGTCTGTGTGTGTCGGGTCTGAGCTCC 1317
QY 1321 CCGGCCAAGCTAGCTCCATGGAAGAAAGAACTTCTCATTCAAAGCAAGAGCTTCAAGAC 1380
Db 1318 CCGGCCAAGCTAGCTCCATGGAAGAAAGAACTTCTCATTCAAAGCAAGAGCTTCAAGAC 1377
QY 1381 TCTCAGGACAAGTGTCAACAAGATGAGACAGGAATGACCCGGTTACATCGGAGAGTCA 1440
Db 1378 TCTCAGGACAAGTGTCAACAAGATGAGACAGGAATGACCCGGTTACATCGGAGAGTCA 1437
QY 1441 GAGGTGAGGCTGTGCTTTAGTCAAGAGAGGTGAGCTGAAGGCTTGAAGCTCAGAGA 1500
Db 1438 GAGGTGAGGCTGTGCTTTAGTCAAGAGAGGTGAGCTGAAGGCTTGAAGCTCAGAGA 1497
QY 1501 TCCCTCCTGAGACAGGACCTTTGCTACCTTACATCAACAATGCAATAGCTTAAAGCGAAT 1560
Db 1498 TCCCTCCTGAGACAGGACCTTTGCTACCTTACATCAACAATGCAATAGCTTAAAGCGAAT 1557
QY 1561 TTGAGACAGCAAGATGAGAGTGTCCAGAGAGATGACAAAGCACTGACGCTTCTCAT 1620
Db 1558 TTGAGACAGCAAGATGAGAGTGTCCAGAGAGATGACAAAGCACTGACGCTTCTCAT 1617
QY 1621 GATATCAGAGACAGAGCCGGAAAGCTTCAAGAAATCAAGAGCAGAGTACAGGCTCAA 1680
Db 1618 GATATCAGAGACAGAGCCGGAAAGCTTCAAGAAATCAAGAGCAGAGTACAGGCTCAA 1677
QY 1681 GTGGAGAAATGAGGTGATGATGAATCAATGTTGGAAGAGATCTTGTCTCAGCAAGAA 1740
Db 1678 GTGGAGAAATGAGGTGATGATGAATCAATGTTGGAAGAGATCTTGTCTCAGCAAGAA 1737
QY 1741 CGGAGTATCTTACGAATCTGAGCTGAGAGAGTCTCGGCTTGCTGCTGAAGATTCAAG 1800
Db 1738 CGGAGTATCTTACGAATCTGAGCTGAGAGAGTCTCGGCTTGCTGCTGAAGATTCAAG 1797
QY 1801 CGGAAGCGACAGAAATGTGAGCATTAACCTGTTGAAGCTTAAGATCAAGGGAAGCCTGAA 1860
Db 1798 CGGAAGCGACAGAAATGTGAGCATTAACCTGTTGAAGCTTAAGATCAAGGGAAGCCTGAA 1857
QY 1861 GTGGGGAATATGCGGAACTGAGAGAAATCAATGCTGAGACAGAGCTCAAAATTCAGAG 1920
Db 1858 GTGGGGAATATGCGGAACTGAGAGAAATCAATGCTGAGACAGAGCTCAAAATTCAGAG 1917
QY 1921 CTCGAAGAGAACTGAGAGAGGCTGTAAAGCCAGACCGGAGGCAACCGAGCTGTGCAG 1980
Db 1918 CTCGAAGAGAACTGAGAGAGGCTGTAAAGCCAGACCGGAGGCAACCGAGCTGTGCAG 1977
QY 1981 AATATCCGCCAGGCAAAAGAGCGAGCCGAGAGGGAGCTGAGAGAGCTGCAGAACCGAGAG 2040
Db 1978 AATATCCGCCAGGCAAAAGAGCGAGCCGAGAGGGAGCTGAGAGAGCTGCAGAACCGAGAG 2037
QY 2041 GATTCTTCTGAAAGGCACTCAGAAAGAGCTGTGTGGAAGCTGAGAGAGCCGCCATTCTCTG 2100
Db 2038 GATTCTTCTGAAAGGCACTCAGAAAGAGCTGTGTGGAAGCTGAGAGAGCCGCCATTCTCTG 2097
QY 2101 GAGAACCAAGGTAAAGAGCTAGAGACCATGAGCCTAGAGAAACAGACTGAAGAGTAGC 2160
Db 2098 GAGAACCAAGGTAAAGAGCTAGAGACCATGAGCCTAGAGAAACAGACTGAAGAGTAGC 2157
QY 2161 ATCCAGACAAAATCCCAACAGATCCAGAGATGAGTGTATTAATTTCTGAGCTCGAAGAG 2220
Db 2158 ATCCAGACAAAATCCCAACAGATCCAGAGATGAGTGTATTAATTTCTGAGCTCGAAGAG 2217
QY 2221 AAAACATCGGAGAGCCCAAGTCTCAGCCAGCACTTGAAGTGCACCTGAACAAGAAAG 2280
Db 2218 AAAACATCGGAGAGCCCAAGTCTCAGCCAGCACTTGAAGTGCACCTGAACAAGAAAG 2277
QY 2281 CAGCACTATGAGAGAAAGATTAAAGTGTGCAATCAATTAAGAAAGAACTGTGCTGAC 2340

2278 CAGCACTATGAGAGAAAGATTAAAGTATTGGAACAATAGATTAAGAAAGACCTGCTGAC 2337
QY 2341 AAGAGACACTGAGAGAAATGATGACAGACAAGAGAGAGAGGCCCATGAGAGGGCAAA 2400
Db 2338 AAGAGACACTGAGAGAAATGATGACAGACAAGAGAGAGAGGCCCATGAGAGGGCAAA 2397
QY 2401 ATTCTACGCCAACAAGAGGGATGATCAATGCTATGGAATTCGAAGATCAATCCCTGGAA 2460
Db 2398 ATTCTACGCCAACAAGAGGGATGATCAATGCTATGGAATTCGAAGATCAATCCCTGGAA 2457
QY 2461 CAGAGATTTGGAATCTGCTGAAGCCAATTAACCTTGCAGCAAAATAGCAGTCTTTTACC 2520
Db 2458 CAGAGATTTGGAATCTGCTGAAGCCAATTAACCTTGCAGCAAAATAGCAGTCTTTTACC 2517
QY 2521 CAAAGAACATGAAGGCCCAAGAGAGATGATTTCTGAACCTCAGGCAACGAAATTTTAC 2580
Db 2518 CAAAGAACATGAAGGCCCAAGAGAGATGATTTCTGAACCTCAGGCAACGAAATTTTAC 2577
QY 2581 CTGAGACACAGGCTGGGAAAGTTGAGAGGCCCAAGAACGAAATCTGAGAGCAGCTGGAG 2640
Db 2578 CTGAGACACAGGCTGGGAAAGTTGAGAGGCCCAAGAACGAAATCTGAGAGCAGCTGGAG 2637
QY 2641 AAGATCAGCCACCAAGACCAAGTGAACAAGATTCGGCTGGAACCTGAGACAAAGATTG 2700
Db 2638 AAGATCAGCCACCAAGACCAAGTGAACAAGATTCGGCTGGAACCTGAGACAAAGATTG 2697
QY 2701 CGGAGGTCAGTGTAGACACAGAGACAGAGAACTGAGCTCAAGGCCAGCTCACAGAG 2760
Db 2698 CGGAGGTCAGTGTAGACACAGAGACAGAGAACTGAGAGCTCAAGGCCAGCTCACAGAG 2757
QY 2761 CTACAGCTCTCCCTGACAGAGCCGAGTCAACATTTGACAGCCTTGCAGAGTTCACCGGCG 2820
Db 2758 CTACAGCTCTCCCTGACAGAGCCGAGTCAACATTTGACAGCCTTGCAGAGTTCACCGGCG 2817
QY 2821 GCCCTGAGAGCCAGCTTGCSCAGGCGAAGACAGAGCTGGAGAGACCAAGCAGAGAGCT 2880
Db 2818 GCCCTGAGAGCCAGCTTGCSCAGGCGAAGACAGAGCTGGAGAGACCAAGCAGAGAGCT 2877
QY 2881 GAAAGAGAGATCCAGGCACTCAAGGCAATAGAGATGAATTCAGGCGCAAAATTTGATGCT 2940
Db 2878 GAAAGAGAGATCCAGGCACTCAAGGCAATAGAGATGAATTCAGGCGCAAAATTTGATGCT 2937
QY 2941 CTTCTGTAACAGCTGTACTGTAAATCAAGACCTGAGAGAGCAGTAAACCAAGCTGACCGAG 3000
Db 2938 CTTCTGTAACAGCTGTACTGTAAATCAAGACCTGAGAGAGCAGTAAACCAAGCTGACCGAG 2997
QY 3001 GACCAAGCTGAAGCTCAACCAACCAAACTTCTGCTGCAAAACAATCTGATGAGGCTTCT 3060
Db 2998 GACCAAGCTGAAGCTCAACCAACCAAACTTCTGCTGCAAAACAATCTGATGAGGCTTCT 3057
QY 3061 GGGGCCAACGACGAGATTGTACAATGCGAAGTGAAGTGAACCATCTCCGCCGGAGATC 3120
Db 3058 GGGGCCAACGACGAGATTGTACAATGCGAAGTGAAGTGAACCATCTCCGCCGGAGATC 3117
QY 3121 ACGGAACGAGATATGAGCTTACCAAGCCAGAGCAAAACGATGAGAGCTTGAAGACCAAG 3180
Db 3118 ACGGAACGAGATATGAGCTTACCAAGCCAGAGCAAAACGATGAGAGCTTGAAGACCAAG 3177
QY 3181 TGCACCAATGCTGAGAGAAACAGGTCATGATTTGAGAGCCCTTAAACGATGAGCTGAGAA 3240
Db 3178 TGCACCAATGCTGAGAGAAACAGGTCATGATTTGAGAGCCCTTAAACGATGAGCTGAGAA 3237
QY 3241 AAAAGAGCCGACAGTGGAGGCTTGAAGAGCGTCTTGGGTGATGAGAAATCCAGTTTGAAG 3300
Db 3238 AAAAGAGCCGACAGTGGAGGCTTGAAGAGCGTCTTGGGTGATGAGAAATCCAGTTTGAAG 3297
QY 3301 TGTGGGTTTCAAGAGCTGCAAGATGCTGCAACCCGAGAAACAAGACAGGCGAGAGGCT 3360
Db 3298 TGTGGGTTTCAAGAGCTGCAAGATGCTGCAACCCGAGAAACAAGACAGGCGAGAGGCT 3357
QY 3361 GATCAGCCGATCAACGAGCTTCCCAAGGTTGAGAGGCTGGCAGTGAAGAGACCAAGGCT 3420
Db 3358 GATCAGCCGATCAACGAGCTTCCCAAGGTTGAGAGGCTGGCAGTGAAGAGACCAAGGCT 3417

Oy	3421	GAGATTCTCGCTCTGCAAGAGCTTCTCAAAGACGAAGCTGAAGGCGCAAGCTCTCT	3480
Db	3418	GAGATTCTCGCTCTGCAAGAGCTTCTCAAAGACGAAGCTGAAGGCGCAAGCTCTCT	3477
Oy	3401	GACAAAGCTCAATGACCTGAGAGAAAGGATGTAATGCTTGAATGAATGCCGAAGCTTA	3540
Db	3478	GACAAAGCTCAATGACCTGAGAGAAAGGATGTAATGCTTGAATGAATGCCGAAGCTTA	3537
Oy	3541	CAGCAGAAGCTGGAGACTGAAAGAGAGCTCAACAGAGGCTTCTGGAAGCAAGCCAA	3600
Db	3538	CAGCAGAAGCTGGAGACTGAAAGAGAGCTCAACAGAGGCTTCTGGAAGCAAGCCAA	3597
Oy	3601	TTACAGCAGCAGATGAGACTTGCAAAAAATCACTATTTCCGTCTGACTCAAGAGACTGCA	3660
Db	3598	TTACAGCAGCAGATGAGACTTGCAAAAAATCACTATTTCCGTCTGACTCAAGAGACTGCA	3657
Oy	3661	GAGGCTCTAGATCGGGCTGATCTACTGTAAGAAGAAAAAGAGTGAAGCTTGAGATCAAGTG	3720
Db	3658	GAGGCTCTAGATCGGGCTGATCTACTGTAAGAAGAAAAAGAGTGAAGCTTGAGATCAAGTG	3717
Oy	3721	GAAGAACATTCAAGTTCTCTATTCTCATGAAAAAGTGAATGGAAGGCACTATTCTCAA	3780
Db	3718	GAAGAACATTCAAGTTCTCTATTCTCATGAAAAAGTGAATGGAAGGCACTATTCTCAA	3777
Oy	3781	CAACCAAACTCATTTGATTTTCTGCAAGCCAAATGACCAACTGCTTAAAAAGAAAAG	3840
Db	3778	CAACCAAACTCATTTGATTTTCTGCAAGCCAAATGACCAACTGCTTAAAAAGAAAAG	3837
Oy	3841	GTTCCCTCTGCAATCAATGAGCTGAAGCTGGCGCTGGAAGAGGAAGAAAGCTCGCTGAGCA	3900
Db	3838	GTTCCCTCTGCAATCAATGAGCTGAAGCTGGCGCTGGAAGAGGAAGAAAGCTCGCTGAGCA	3897
Oy	3901	GAGCTTAGAGAAAGCCCTTCAAGAAAGCCGATCGAGCTCGGCTCGCGCGGAGAGAAAGCT	3960
Db	3898	GAGCTTAGAGAAAGCCCTTCAAGAAAGCCGATCGAGCTCGGCTCGCGCGGAGAGAAAGCT	3957
Oy	3961	GCCCAACCGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG	4020
Db	3958	GCCCAACCGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG	4017
Oy	4021	ATGCGCATGTCGCGCATGTCGCGATGCGCAGAGCACAGCCAGTGCATGAGCTGCTG	4080
Db	4018	ATGCGCATGTCGCGCATGTCGCGATGCGCAGAGCACAGCCAGTGCATGAGCTGCTG	4077
Oy	4081	GCCCCGCGCATCCAGCCGAGAAAGAGTCTTCAACTCCAGAGGAATTTAGTCGGCGCTTT	4140
Db	4078	GCCCCGCGCATCCAGCCGAGAAAGAGTCTTCAACTCCAGAGGAATTTAGTCGGCGCTTT	4137
Oy	4141	AAGGAAGCGATGACACCACTATTTCTCAACGGATTTCAAGTGAAGTCAACATGCAAGCC	4200
Db	4138	AAGGAAGCGATGACACCACTATTTCTCAACGGATTTCAAGTGAAGTCAACATGCAAGCC	4197
Oy	4201	ACAAAGTGTGCTGTGTCTGGATACGCTGACTTTGACGCGCAGAGGATCCAAATGCTCTC	4260
Db	4198	ACAAAGTGTGCTGTGTCTGGATACGCTGACTTTGACGCGCAGAGGATCCAAATGCTCTC	4257
Oy	4261	GAATGTCAAGTGTGTGTCAACCCCAAGTCTTCAAGTGTCTTCCAGCAGCACTTGCGCTTG	4320
Db	4258	GAATGTCAAGTGTGTGTCAACCCCAAGTCTTCAAGTGTCTTCCAGCAGCACTTGCGCTTG	4317
Oy	4321	CCTGCTGAATATGCAACACTTTCACGAGGCTTCTGCGCTGACAAATGAATCACTCCCCA	4380
Db	4318	CCTGCTGAATATGCAACACTTTCACGAGGCTTCTGCGCTGACAAATGAATCACTCCCCA	4377
Oy	4381	GATCTCCAGACCAAGAGCCCGACAGACTTGCACTGGAAGGCTGATGAAGTGTGCC	4440
Db	4378	GATCTCCAGACCAAGAGCCCGACAGACTTGCACTGGAAGGCTGATGAAGTGTGCC	4437
Oy	4441	AGGAATTAACAAACGAGGACAGCAAGGCTGGGACAGAAAGTACATTGTCTTGAGAGGATCA	4500
Db	4438	AGGAATTAACAAACGAGGACAGCAAGGCTGGGACAGAAAGTACATTGTCTTGAGAGGATCA	4497

QY	4501	AAAGTCCTCATTATATGACATGAAAGCCAGAGAACTGACAGAGCCGGTGTGAAAGATT	4560
Db	4498	AAAGTCCTCATTATATGACATGAAAGCCAGAGAACTGACAGAGCCGGTGTGAAAGATT	4557
QY	4551	GAGCTGTGCTTCCCGACGGGGAGTATCTATTCATGATGGTCCGTTGGTCTCCGAATC	4620
Db	4558	GAGCTGTGCTTCCCGACGGGGAGTATCTATTCATGATGGTCCGTTGGTCTCCGAATC	4617
QY	4621	GCAATACAGCCAAAGCAGATGTCCCATACATACAGATGAAATCTCACCCGACACC	4680
Db	4618	GCAATACAGCCAAAGCAGATGTCCCATACATACAGATGAAATCTCACCCGACACC	4677
QY	4661	AACCTGTGCCCCGGGAGAACTCTTACTGTACTCCAGCTTCCCTGACAAACAGCGC	4740
Db	4678	AACCTGTGCCCCGGGAGAACTCTTACTGTACTCCAGCTTCCCTGACAAACAGCGC	4737
QY	4741	TGGGTACCCGCTTATGAAATCAGTGTGTGGCAGGTGGGAGAGTTTGTAGGGAAAAAGCAGA	4800
Db	4738	TGGGTACCCGCTTATGAAATCAGTGTGTGGCAGGTGGGAGAGTTTGTAGGGAAAAAGCAGA	4797
QY	4801	GCTGATGCTAAATCGTTTGAAACTCCCTGCTGAACCTGMACTGMAAGGTGATGACCGTCTAGAC	4860
Db	4798	GCTGATGCTAAATCGTTTGAAACTCCCTGCTGAACCTGMAAGGTGATGACCGTCTAGAC	4857
QY	4861	ATGAATCTGCACGCTGCCCTTCAGTGACACAGGTGTGTTGGTGGACCCGAGAAAGGCTC	4920
Db	4858	ATGAATCTGCACGCTGCCCTTCAGTGACACAGGTGTGTTGGTGGACCCGAGAAAGGCTC	4917
QY	4921	TACGCCCTGAATGTCTTGAAAAATCTCCCTAATCCCATGCTCCAGGAATTGGACAGCTTTC	4980
Db	4918	TACGCCCTGAATGTCTTGAAAAATCTCCCTAATCCCATGCTCCAGGAATTGGACAGCTTTC	4977
QY	4981	CAAAATTTATATATCAAGAGCTGTGAGAAAGCTACATGATAGCAGAGAAAGACGGGACA	5040
Db	4978	CAAAATTTATATATCAAGAGCTGTGAGAAAGCTACATGATAGCAGAGAAAGACGGGACA	5037
QY	5041	CTGTGTCTTGTGACGTGAGAAAGTGAAACAGTCCCTGGCCAGTCCCACTCTGCTGCC	5100
Db	5038	CTGTGTCTTGTGACGTGAGAAAGTGAAACAGTCCCTGGCCAGTCCCACTCTGCTGCC	5097
QY	5101	CAGCCCGACATCTCACCCCAATTTTGAAGCTGTCAAGGGCTGSCCATTTGTGGGGACA	5160
Db	5098	CAGCCCGACATCTCACCCCAATTTTGAAGCTGTCAAGGGCTGSCCATTTGTGGGGACA	5157
QY	5161	GGAAGAATTGAAACGGGCTCTGCATCTGTGACACCATGCCCCAGCAAAAGTCTGATTTCTC	5220
Db	5158	GGAAGAATTGAAACGGGCTCTGCATCTGTGACACCATGCCCCAGCAAAAGTCTGATTTCTC	5217
QY	5221	CGCTACACGAAAACTTCAGCAAAATCTGCATCTCCGAGAAAGATAGAGACTCTAGAGCCC	5280
Db	5218	CGCTACACGAAAACTTCAGCAAAATCTGCATCTCCGAGAAAGATAGAGACTCTAGAGCCC	5277
QY	5281	TGCAGCTGTATCCCACTTTCACCAATTAACGATATCTCATTTGGAACCAATTAATTCAGAA	5340
Db	5278	TGCAGCTGTATCCCACTTTCACCAATTAACGATATCTCATTTGGAACCAATTAATTCAGAA	5337
QY	5341	ATGCACATGAGACGATACACGCTCGAGGAATTTCTGTGATAGAAATGACATCTCTTGACA	5400
Db	5338	ATGCACATGAGACGATACACGCTCGAGGAATTTCTGTGATAGAAATGACATCTCTTGACA	5397
QY	5401	CCTGTGTGTGTGCGCCTCTTCCAAACGCTTCCCTGTTCATCTGTGACAGTGAACAGC	5460
Db	5398	CCTGTGTGTGTGCGCCTCTTCCAAACGCTTCCCTGTTCATCTGTGACAGTGAACAGC	5457
QY	5461	GCAGGGCAGCGAAGAGATATCTTCTGTGTTCACGAAATTTGGAAGTTCGTGGATTTCT	5520
Db	5458	GCAGGGCAGCGAAGAGATATCTTCTGTGTTCACGAAATTTGGAAGTTCGTGGATTTCT	5517
QY	5521	TACGGAAGACGTACCCGACAGAGATCTCAAGTGAAGTGCCTTAACCTTTGGCTTTGGC	5580
Db	5518	TACGGAAGACGTACCCGACAGAGATCTCAAGTGAAGTGCCTTAACCTTTGGCTTTGGC	5577
QY	5581	TACAGAGAACCTTATCTGTGTTGACCACTTCACTCATCGAAGTAATTTGAGATCCAG	5640

Db 5578 TACAGAAACCCATCTCTGTTTGTATCCACTTCACTCAAGTAATTAATTCAG 5637
Qy 5641 GCACGCTCTCAGAGAGAGACCCCTGCCGAGCGTACCTGACATCCCGAACCCGCGCTAC 5700
Db 5638 GCACGCTCTCAGAGAGAGACCCCTGCCGAGCGTACCTGACATCCCGAACCCGCGCTAC 5697
Qy 5701 CTGGGCCCCGCAATTTCTCTCAGAGAGCAATTAATTCTGGCGTCTCATACAGATTAATTA 5760
Db 5698 CTGGGCCCCGCAATTTCTCTCAGAGAGCAATTAATTCTGGCGTCTCATACAGATTAATTA 5757
Qy 5761 AGGGTCAATTTGCTCAGAGAGAACTCTGTGAAGAGTCCGGCACTGAACACCAACCGGGGC 5820
Db 5758 AGGGTCAATTTGCTCAGAGAGAACTCTGTGAAGAGTCCGGCACTGAACACCAACCGGGGC 5817
Qy 5821 CCGTCCACTCTCCGAG 5880
Db 5818 CCGTCCACTCTCCGAG 5877
Qy 5881 ACCAAGCGCTGTGCTCTCAG 5940
Db 5878 ACCAAGCGCTGTGCTCTCAG 5937
Qy 5941 AGCACAACCCACCGCTACCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6000
Db 5938 AGCACAACCCACCGCTACCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5997
Qy 6001 CGCCTCCGAG 6060
Db 5998 CGCCTCCGAG 6057
Qy 6061 CCGGCGAG 6120
Db 6058 CCGGCGAG 6117
Qy 6121 CCGCTGTCCAG 6165
Db 6118 CCGCTGTCCAG 6162

RESULT 10
ADN62806
ID ADN62806 standard; DNA; 6189 BP.
AC ADN62806;
XX
XX 01-JUL-2004 (first entry)
XX
XX Human NOV1a DNA.
XX
XX ds; gene; human; NOVX; metabolic disorder; diabetes; obesity;
XX infectious disease; anorexia; cancer; cancer-associated cachexia;
XX neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;
XX immune disorder; haematopoietic disorder; dyslipidaemia;
XX metabolic syndrome X; wasting disorder; SNP;
XX single nucleotide polymorphism.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX variation replace(4319,t)
XX FT /tag= a
XX FT /note= "Single nucleotide polymorphism"
XX
XX US2004038223-A1.
XX
XX 26-FEB-2004.
XX
XX 01-OCT-2002; 2002US-00262511.
XX
XX 02-OCT-2001; 2001US-0326483P.
XX 05-OCT-2001; 2001US-0327435P.
XX 05-OCT-2001; 2001US-0327449P.
XX
XX

PR 09-OCT-2001; 2001US-0327917P.
PR 09-OCT-2001; 2001US-0328029P.
PR 09-OCT-2001; 2001US-0328044P.
PR 09-OCT-2001; 2001US-0328056P.
PR 12-OCT-2001; 2001US-0328849P.
PR 15-OCT-2001; 2001US-0329414P.
PR 17-OCT-2001; 2001US-0330142P.
PR 18-OCT-2001; 2001US-0330309P.
PR 22-OCT-2001; 2001US-0341058P.
PR 24-OCT-2001; 2001US-0339266P.
PR 24-OCT-2001; 2001US-0345629P.
PR 29-OCT-2001; 2001US-0349575P.
PR 01-NOV-2001; 2001US-0346357P.
PR 17-APR-2002; 2002US-0373260P.
PR 19-APR-2002; 2002US-0373815P.
PR 19-APR-2002; 2002US-0373817P.
PR 19-APR-2002; 2002US-0373826P.
PR 19-APR-2002; 2002US-0373884P.
PR 22-APR-2002; 2002US-0374977P.
PR 16-MAY-2002; 2002US-0381037P.
PR 16-MAY-2002; 2002US-0381038P.
PR 16-MAY-2002; 2002US-0381042P.
PR 17-MAY-2002; 2002US-0381642P.
PR 28-MAY-2002; 2002US-0383566P.
PR 29-MAY-2002; 2002US-0383831P.
PR 25-JUN-2002; 2002US-0391335P.
XX
PA (SMIT/) SMITHSON G.
PA (MILL/) MILLET I.
PA (PEYM/) PEYMAN J A.
PA (KEKU/) KEKUDA R.
PA (UJUU/) JU J.
PA (LILL/) LI L.
PA (GUOX/) GUO X.
PA (BATT/) PATTURAJAN M.
PA (SPYT/) SPYTEK K A.
PA (EDIN/) EDINGER S R.
PA (ELLE/) ELLERMAN K.
PA (MALY/) MALYANKAR U M.
PA (ORTT/) ORT T.
PA (GORM/) GORMAN L.
PA (ZERR/) ZERTUSEN B D.
PA (ANDE/) ANDERSON D W.
PA (ZHON/) ZHONG M.
PA (CATT/) CATTERTON E.
PA (JIMW/) JI W.
PA (MILL/) MILLER C E.
PA (RAST/) RASTELLI L.
PA (STON/) STONE D J.
PA (PENNA/) PENNA C E A.
PA (SHEN/) SHENOY S G.
PA (SHIM/) SHINKERS R A.
PA (ROTH/) ROTHENBERG M E.
PA (LEAC/) LEACH M D.
PA (AGEE/) AGEE M L.
PA (BERG/) BERGHS C.
PA (DIP/) DIPIPPO V A.
PA (EISE/) EISEN A.
PA (GANG/) GANGOLLI E A.
PA (RIEG/) RIEGER D K.
PA (SPAD/) SPADERNA S K.
XX
PI Smithson G, Millet I, Peyman JA, Kekuda R, Ju J, Li L, Guo X;
PI Paturajan M, Spytke KA, Edinger SR, Ellerman K, Malynkar UM;
PI Ort T, Gorman L, Zertusen BD, Anderson DW, Zhong M, Catterton E;
PI Ji W, Miller CE, Rastelli L, Stone DJ, Pena CE, Shenoy SG;
PI Shinkers RA, Rothenberg ME, Leach MD, Agee ML, Bergs C, Dipippo VA;
PI Eissen A, Gangolli EA, Rieger DK, Spaderna SK;
XX
XX MPI; 2004-213931/20.
XX P-PsDB; ADN62807.
XX
XX Isolated NOVX polypeptides and nucleic acids, useful for preventing,
XX


```
QY 1621 GATATCAGAGCGAGACCCGAAAGCTCCAGAAATCAAGAAGCGAGATACAGAGCTCCTAA 1680
DB 1618 GATATCAGAGCGAGACCCGAAAGCTCCAGAAATCAAGAAGCGAGATACAGAGCTCCTAA 1677
QY 1681 GTGGAAAGAAATGAGGTGATGATGAATCAGTTGGAGAGGATCTTGTCTCGACAAAGAA 1740
DB 1678 GTGGAAAGAAATGAGGTGATGATGAATCAGTTGGAGAGGATCTTGTCTCGACAAAGAA 1737
QY 1741 CGGAGTGATCTTACGAAATCTGAGCTGAGAGAGTCTCGGCTTGCTGTAAGAAATTCAG 1800
DB 1738 CGGAGTGATCTTACGAAATCTGAGCTGAGAGAGTCTCGGCTTGCTGTAAGAAATTCAG 1797
QY 1801 CGGAAGCGCAGAAATCTCAGCAATTAACCTGTTGAAGGCTTAAGATCAAGGAAAGCTCGAA 1860
DB 1798 CGGAAGCGCAGAAATCTCAGCAATTAACCTGTTGAAGGCTTAAGATCAAGGAAAGCTCGAA 1857
QY 1861 GTGGAGAAATATGCGAAATCTGAGAAAGATCAATGCTGAGCAGCACTCAAAATTCAGAG 1920
DB 1858 GTGGAGAAATATGCGAAATCTGAGAAAGATCAATGCTGAGCAGCACTCAAAATTCAGAG 1917
QY 1921 CTCCAAGAGAAATCTGAGAAAGGCTGTAAAGCCAGCAGCGAGGCACTCGAGCTGCTGAG 1980
DB 1918 CTCCAAGAGAAATCTGAGAAAGGCTGTAAAGCCAGCAGCGAGGCACTCGAGCTGCTGAG 1977
QY 1981 AATATCCGCGAGGCAAAAGAGCGGCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2040
DB 1978 AATATCCGCGAGGCAAAAGAGCGGCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2037
QY 2041 GATTCTTCTGAAAGGCAATCAGAAAGAGCTGTGAAAGCTGAGAAAGCCGCGCATTCCTCTG 2100
DB 2038 GATTCTTCTGAAAGGCAATCAGAAAGAGCTGTGAAAGCTGAGAAAGCCGCGCATTCCTCTG 2097
QY 2101 GAGAACCAAGGTAAAGAGCTAGAGACCATGAGCGTAGAGAAACAGACTGAAGAGATGAC 2160
DB 2098 GAGAACCAAGGTAAAGAGCTAGAGACCATGAGCGTAGAGAAACAGACTGAAGAGATGAC 2157
QY 2161 ATCCAGACCAAAATCCCAACAGATCCAGCAGATGCTGATTAATTTCTGAGCTCGAAGAG 2220
DB 2158 ATCCAGACCAAAATCCCAACAGATCCAGCAGATGCTGATTAATTTCTGAGCTCGAAGAG 2217
QY 2221 AAAACATCGGAGGCGCCAGAGTCTGAGCGCAGCAGCTTGAAGTGCACCTGAAACGAAAGAG 2280
DB 2218 AAAACATCGGAGGCGCCAGAGTCTGAGCGCAGCAGCTTGAAGTGCACCTGAAACGAAAGAG 2277
QY 2281 CAGCACTATGAGAAAGATTAAAGTGTGACCAATCAGATTAAGAAAGACTGCGCTGAC 2340
DB 2278 CAGCACTATGAGAAAGATTAAAGTGTGACCAATCAGATTAAGAAAGACTGCGCTGAC 2337
QY 2341 AAGGAGACACTGAGAAACATGATGCAAGACAGAGAGAGGCGCATGAGAAAGGCGCAA 2400
DB 2338 AAGGAGACACTGAGAAACATGATGCAAGACAGAGAGAGGCGCATGAGAAAGGCGCAA 2397
QY 2401 ATTCTCAGCGAAACAGAAAGCGATGATCAATGCTATGGAATTCGAAGTCAAGTCCCTGGA 2460
DB 2398 ATTCTCAGCGAAACAGAAAGCGATGATCAATGCTATGGAATTCGAAGTCAAGTCCCTGGA 2457
QY 2461 CAGAGATTTGGAACCTGCTGAAGCCAAATTAACCTTGACAGCAATGACATCTTTTACC 2520
DB 2458 CAGAGATTTGGAACCTGCTGAAGCCAAATTAACCTTGACAGCAATGACATCTTTTACC 2517
QY 2521 CAAAGGACATGAAAGGCCCAAGAGAGATGATTTCTGAACCTCAGCAACAGAAATTTTAC 2580
DB 2518 CAAAGGACATGAAAGGCCCAAGAGAGATGATTTCTGAACCTCAGCAACAGAAATTTTAC 2577
QY 2581 CTGGAGACACAGGCTGGGGAAGTTGAGGCGCCAGAAACCCGAAACCTGAGAGAGAGCTGGAG 2640
DB 2578 CTGGAGACACAGGCTGGGGAAGTTGAGGCGCCAGAAACCCGAAACCTGAGAGAGAGCTGGAG 2637
QY 2641 AAGATCAGCCACCAAGACCAAGTGAACAAGAAATCGGCTCTGGAATCTGAGACCAAGATTG 2700
DB 2638 AAGATCAGCCACCAAGACCAAGTGAACAAGAAATCGGCTCTGGAATCTGAGACCAAGATTG 2697
QY 2701 CGGAGGTCAGTCTAGACGACAGAGAGCAGAAACCTGAGGCTCAAGCCCAAGCTCAAGAG 2760
DB 2698 CGGAGGTCAGTCTAGACGACAGAGAGCAGAAACCTGAGGCTCAAGCCCAAGCTCAAGAG 2757
QY 2761 CTACAGCTCCCTCGACGAGCGCGAGTCAAGTTGACAGCCCTCGAGGCTGACCGGCGG 2820
DB 2758 CTACAGCTCCCTCGACGAGCGCGAGTCAAGTTGACAGCCCTCGAGGCTGACCGGCGG 2817
QY 2821 GCCCTGAGAGCCAGCTTCGCGAGGCGAGAGCAGAGCTGGAAGAGACCAAGCAGAGGCT 2880
DB 2818 GCCCTGAGAGCCAGCTTCGCGAGGCGAGAGCAGAGCTGGAAGAGACCAAGCAGAGGCT 2877
QY 2881 GAAGAGAGATCCAGGCACTCAGCGACATGAGATGAATTCAGCGCAAAATTTGATGCT 2940
DB 2878 GAAGAGAGATCCAGGCACTCAGCGACATGAGATGAATTCAGCGCAAAATTTGATGCT 2937
QY 2941 CTTCTGTAACAGCTGTACTGTATCAAGACCTTGAAGAGAGGCTAAACAGCTGACCGAG 3000
DB 2938 CTTCTGTAACAGCTGTACTGTATCAAGACCTTGAAGAGAGGCTAAACAGCTGACCGAG 2997
QY 3001 GACAAAGCTGAACTCAACAAACCAAACTTCTACTGTCCAAAACCTGATGAGGCTTCT 3060
DB 2998 GACAAAGCTGAACTCAACAAACCAAACTTCTACTGTCCAAAACCTGATGAGGCTTCT 3057
QY 3061 GCGCGCAACGACGAGATTGTAACAATGCGAAGTGAAGTGAACCATTCGCGCGAGATC 3120
DB 3058 GCGCGCAACGACGAGATTGTAACAATGCGAAGTGAAGTGAACCATTCGCGCGAGATC 3117
QY 3121 ACGGAACGAGATGAGAGCTTACAGCGAAGCAACGATGGAAGGCTGGAAGACCAAG 3180
DB 3118 ACGGAACGAGATGAGAGCTTACAGCGAAGCAACGATGGAAGGCTGGAAGACCAAG 3177
QY 3181 TGCACCATGCTGAGAGAAAGATCATGATTTGGAAGGCGCTTAACCATGAGCTGTAGAA 3240
DB 3178 TGCACCATGCTGAGAGAAAGATCATGATTTGGAAGGCGCTTAACCATGAGCTGTAGAA 3237
QY 3241 AAAGAGCGCAGTGGGAGGCTGAGAGAGGCTCTGGGTGATGAGAAATCCAGTTTGA 3300
DB 3238 AAAGAGCGCAGTGGGAGGCTGAGAGAGGCTCTGGGTGATGAGAAATCCAGTTTGA 3297
QY 3301 TGTGGGTTCTGAGAGCTGAGAGGATGCTGAGACCGCAAGCAAGAGCAGGCGGAGAGGC 3360
DB 3298 TGTGGGTTCTGAGAGCTGAGAGGATGCTGAGACCGCAAGCAAGAGCAGGCGGAGAGGC 3357
QY 3361 GATCAGCGGATCACCAAGTCTCGCCAGGTGTGAGCTGCAAGTGAAGAGCAAGAGGCT 3420
DB 3358 GATCAGCGGATCACCAAGTCTCGCCAGGTGTGAGCTGCAAGTGAAGAGCAAGAGGCT 3417
QY 3421 GAGATTTCTGCTGTGACAGAGGCTCTCAAAAGCAGAAAGCTGAAGGCGAGGCTTCT 3480
DB 3418 GAGATTTCTGCTGTGACAGAGGCTCTCAAAAGCAGAAAGCTGAAGGCGAGGCTTCT 3477
QY 3481 GACAAAGCTCAATGACCTGGAGAAAGAGATGCTATGAAATGAAATGACCCGAGGCTTA 3540
DB 3478 GACAAAGCTCAATGACCTGGAGAAAGAGATGCTATGAAATGAAATGACCCGAGGCTTA 3537
QY 3541 CAGCAAGAGCTGAGACTGAAAGAGGCTCAACAGAGGCTTCTGGAAGAGCAAGCCAAA 3600
DB 3538 CAGCAAGAGCTGAGACTGAAAGAGGCTCAACAGAGGCTTCTGGAAGAGCAAGCCAAA 3597
QY 3601 TTACAGCAGAGATGAGCTGCAAGAAATCAATTTTCGTTGACTCAAGAGCTGCAA 3660
DB 3598 TTACAGCAGAGATGAGCTGCAAGAAATCAATTTTCGTTGACTCAAGAGCTGCAA 3657
QY 3661 GAAAGCTTGAATCGGCGTGAATCTTAAGAGAAAGAAAGAGAGCTTGAAGATACAGCTG 3720
DB 3658 GAAAGCTTGAATCGGCGTGAATCTTAAGAGAAAGAAAGAGAGCTTGAAGATACAGCTG 3717
QY 3721 GAAACATTTCAAGGTTCTATTTCTATGAAAGGTAAGTGAAGGCACTATTTCTCAA 3780
DB 3718 GAAACATTTCAAGGTTCTATTTCTATGAAAGGTAAGTGAAGGCACTATTTCTCAA 3777
QY 3781 CAAACCAAACTCATTTGATTTTCTGACAGCCAAATGAGCAACCTGTAAAGAAAGAAAG 3840
```

|||||
Db 3778 CAAACCAACTCATGATTTTTCGCAAGCCAAAATGACCAACTGTCTAAAAAGAAAAAG 3837
Qy 3841 GTTCTCTGAGTACAAATGAGCTGAAGCTGGCCCTGGAGAGAGAAAGCTCGCTGTGCA 3900
Db 3838 GTGCTCTGAGTACAAATGAGCTGAAGCTGGCCCTGGAGAGAGAAAGCTCGCTGTGCA 3897
Qy 3901 GAGCTAGAGAGAGCCCTTCAGAAAGCCCGCATCGAGCTCGGCTCGCCGAGAGAGAGCT 3960
Db 3898 GAGCTAGAGAGAGCCCTTCAGAAAGCCCGCATCGAGCTCGGCTCGCCGAGAGAGAGCT 3957
Qy 3961 GCCCACCAG 4020
Db 3958 GCCCACCAG 4017
Qy 4021 ATGCGCATGTCCGCTCATGTCGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 4080
Db 4018 ATGCGCATGTCTGCTCATGTCGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 4077
Qy 4081 GCCCGCGCATCGAGCCGAG 4140
Db 4078 GCCCGCGCATCGAGCCGAG 4137
Qy 4141 AAGGAAGCATGCAACCAATATTTCTCACCAGATTCAAGTGAAGTGAACATGCGAGCC 4200
Db 4138 AAGGAAGCATGCAACCAATATTTCTCACCAGATTCAAGTGAAGTGAACATGCGAGCC 4197
Qy 4201 ACAAGGT 4260
Db 4198 ACAAGGT 4257
Qy 4261 GAATGTGAGT 4320
Db 4258 GAATGTGAGT 4317
Qy 4321 CTGTGTGAATGT 4380
Db 4318 CTGTGTGAATGT 4377
Qy 4381 GGTCTTCAGACCAAG 4440
Db 4378 GGTCTTCAGACCAAG 4437
Qy 4441 AGGAATTAACAAAG 4500
Db 4438 AGGAATTAACAAAG 4497
Qy 4501 AAAGTCTCTATTATGACAAATGAAGCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4560
Db 4498 AAAGTCTCTATTATGACAAATGAAGCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4557
Qy 4561 GAGGT 4620
Db 4558 GAGGT 4617
Qy 4621 GCAAATTAACAGCAAG 4680
Db 4618 GCAAATTAACAGCAAG 4677
Qy 4681 ACCTGT 4740
Db 4678 ACCTGT 4737
Qy 4741 TGGGTGACCGGCTTTAAGATCAAGT 4800
Db 4738 TGGGTGACCGGCTTTAAGATCAAGT 4797
Qy 4801 GGTGATGCTAAAGCTTTGTGGAACCTCCGTGTGAAATCTGGAAGGTGATGACCGGTCTGAGC 4860
Db 4798 GGTGATGCTAAAGCTTTGTGGAACCTCCGTGTGAAATCTGGAAGGTGATGACCGGTCTGAGC 4857
Qy 4861 ATGAACTGACGCTGT 4920

Db 4858 ATGAACTGACGCTGT 4917
Qy 4921 TACGCCCTGAATGTCTTGAAGAAATCTCCTTAACCATGTCTCCAGAGATTTGAGACGTCTTC 4980
Db 4918 TACGCCCTGAATGTCTTGAAGAAATCTCCTTAACCATGTCTCCAGAGATTTGAGACGTCTTC 4977
Qy 4981 CAAATTTATTTATTTAAGAGCTGTGAG 5040
Db 4978 CAAATTTATTTATTTAAGAGCTGTGAG 5037
Qy 5041 CTGT 5100
Db 5038 CTGT 5097
Qy 5101 CAGCCGAGACATCTCACCACCAATTTTGAAGCTGTCAAGGCTGTGCCACTTGTGTGTGTGTGTGTGTGTGT 5160
Db 5098 CAGCCGAGACATCTCACCACCAATTTTGAAGCTGTCAAGGCTGTGCCACTTGTGTGTGTGTGTGTGTGTGT 5157
Qy 5161 GCGAAGTTGAGAAAGGAGCTGT 5220
Db 5158 GCGAAGTTGAGAAAGGAGCTGT 5217
Qy 5221 CGCTACAGAGAAACCTCAGCAAAATCTGATCCGAGAAAGATGAGAGACTCAGAGCCG 5280
Db 5218 CGCTACAGAGAAACCTCAGCAAAATCTGATCCGAGAAAGATGAGAGACTCAGAGCCG 5277
Qy 5281 TGCAAGCTGTATCCACTTCAACCAATTAACATGATCTCTAATGGAACCAATTAATTTCTAGAA 5340
Db 5278 TGCAAGCTGTATCCACTTCAACCAATTAACATGATCTCTAATGGAACCAATTAATTTCTAGAA 5337
Qy 5341 ATGCAATGAGAGAGATCAAGCTGTGAGAGATTTCTGTGATTAAGATGCAATTTCTTTGCA 5400
Db 5338 ATGCAATGAGAGAGATCAAGCTGTGAGAGATTTCTGTGATTAAGATGCAATTTCTTTGCA 5397
Qy 5401 CTGT 5460
Db 5398 CTGT 5457
Qy 5461 GCAGGAG 5520
Db 5458 GCAGGAG 5517
Qy 5521 TACGAAAGAGATGAG 5580
Db 5518 TACGAAAGAGATGAG 5577
Qy 5581 TACGAGAAACCTTATGT 5640
Db 5578 TACGAGAAACCTTATGT 5637
Qy 5641 GCAAGCTCTCAG 5700
Db 5638 GCAAGCTCTCAG 5697
Qy 5701 CTGGGCGCTGT 5760
Db 5698 CTGGGCGCTGT 5757
Qy 5761 AGGTGATTTGT 5820
Db 5758 AGGTGATTTGT 5817
Qy 5821 CCGTCCACCTCCGAG 5880
Db 5818 CCGTCCACCTCCGAG 5877
Qy 5881 ACCAAGCGCTGT 5940
Db 5878 ACCAAGCGCTGT 5937
Qy 5941 AGCAACCCCAACGCTTACCGGAG 6000
Db 5938 AGCAACCCCAACGCTTACCGGAG 5997

QY 6001 CGCCCCCTGAGCAGAGAGAGTCCCCCGCGGATCTACAGACCGGAGAGACCGGTCC 6060
 DB 5998 CGCCCCCTGAGCAGAGAGAGTCCCCCGCGGATCTACAGACCGGAGAGACCGGTCC 6057
 QY 6061 CCCGCGAGGCTGTTTGAAGACAGACAGAGGGGCGGCTGCTGGCGGAGCCGTGAGGACC 6120
 DB 6058 CCGCGGAGGCTGTTTGAAGACAGACAGAGGGGCGGCTGCTGGCGGAGCCGTGAGGACC 6117
 QY 6121 CCGCTGTCCAGAGTGAACAAGTCTGGAGACCAAGTCTTCAATATA 6165
 DB 6118 CCGCTGTCCAGAGTGAACAAGTCTGGAGACCAAGTCTTCAATATA 6162

RESULT 11
 ABS63435
 ID ABS63435 standard; cDNA; 6201 BP.
 AC ABS63435;
 XX
 DT 15-NOV-2002 (first entry)
 XX
 DE Human cDNA, homologous to kinases, designated NOV3a.
 XX
 KM Human; gene; ss; NOV; NOVX; NOVX-associated disorder; cardiomyopathy;
 KM atherosclerosis; diabetes; cell signalling; metabolic pathway;
 KM cellular receptor; downstream effector; cancer; gene therapy;
 KM hypertension; congenital heart defect; aortic stenosis; obesity;
 KM infectious disease; anorexia; Alzheimer's disease; Parkinson's disease;
 KM neurodegenerative disorder; haemophilia; dyslipidemia; vaccine;
 KM haematopoietic disease; scleroderma; fertility; immunogen;
 KM idiopathic thrombocytopenic purpura; graft versus host disease;
 KM Crohn's disease; multiple sclerosis; cirrhosis; autoimmune disease;
 KM systemic lupus erythematosus; asthma; arthritis; psoriasis; allergy;
 KM stroke; anxiety; Lesch-Nyhan syndrome; schizophrenia; cerebellar ataxia;
 KM pain; alcoholism; transgenic.
 XX
 OS Homo sapiens.
 XX
 PN WO200226826-A2.
 XX
 PD 04-APR-2002.
 XX
 PF 27-SEP-2001; 2001WO-US042336.
 XX
 PR 27-SEP-2000; 2000US-0235631P.
 XX
 PR 27-SEP-2000; 2000US-0235633P.
 PR 27-SEP-2000; 2000US-0235808P.
 PR 27-SEP-2000; 2000US-0236064P.
 PR 27-SEP-2000; 2000US-0236065P.
 PR 27-SEP-2000; 2000US-0236066P.
 PR 28-SEP-2000; 2000US-0236135P.
 PR 03-OCT-2000; 2000US-0237434P.
 PR 05-OCT-2000; 2000US-0238321P.
 PR 06-OCT-2000; 2000US-0238396P.
 PR 06-OCT-2000; 2000US-0238399P.
 PR 16-MAR-2001; 2001US-0276676P.
 PR 31-MAY-2001; 2001US-0294823P.
 PR 12-JUL-2001; 2001US-0304688P.
 PR 26-SEP-2001; 2001US-00235631.
 XX
 PA (CUBA-) CUBAGEN CORP.
 XX
 PI Gerlach VL, MacDougall JR, Smitheon G, Millet I, Stone D;
 PI Gunther E, Ellerman K, Grosse WM, Alsobrook JP, Lepley DM;
 PI Burgess CE, Padigar M, Kekuda R, Spytek KA, Leach MD, Shimkets RA;
 XX
 XX MPI; 2002-499860/53.
 DR P-PSDB; ABG78362.
 DR
 XX
 PT Novel isolated NOVX polypeptides and polynucleotides homologous to
 PT attraction, plexin, plexin-like family of proteins, useful for treating
 PT atherosclerosis, diabetes, cancer, Alzheimer's disease, hemophilia and

PT stroke.
 XX
 PS Claim 8; Page 40-42; 308pp; English.
 XX
 CC The invention discloses the isolated human polypeptides, and
 CC polynucleotides encoding them, that have been designated NOVX. The
 CC polypeptides, polynucleotides and antibodies are useful in treating or
 CC preventing a NOVX-associated disorder which is cardiomyopathy,
 CC atherosclerosis and diabetes in a human, where the disorder is related to
 CC cell signal processing and metabolic pathway modulation. They can also be
 CC used in determining the presence of, or predisposition to, a disease
 CC associated with altered levels of the polypeptides and polynucleotides of
 CC any one of the 13 sequences (NOV1-NOV8), for raising antibodies, for
 CC identifying an agent that binds to, or that modulates the expression or
 CC activity of the polypeptide, for identifying an agent which is cellular
 CC receptor or downstream effector, for treating or preventing a NOVX-
 CC associated disorder and as a pharmaceutical composition comprising the
 CC polypeptide, polynucleotide or the antibody. The polypeptides and
 CC polynucleotides are useful in diagnostic applications (e.g. as a marker
 CC for cancerous cells or tissue types) where their amounts are assessed, or
 CC for the manufacture of a medicament (e.g. gene therapy) for treating or
 CC preventing disorders or syndromes such as hypertension, congenital heart
 CC defects, aortic stenosis, obesity, infectious disease, anorexia, cancer,
 CC Alzheimer's disease, Parkinson's disorder, neurodegenerative disorders,
 CC haemophilia, dyslipidemia, haematopoietic diseases, scleroderma,
 CC fertility, idiopathic thrombocytopenic purpura, graft versus host
 CC diseases, Crohn's disease, multiple sclerosis, cirrhosis, autoimmune
 CC disease, systemic lupus erythematosus, asthma, arthritis, psoriasis,
 CC allergy, stroke, anxiety, Lesch-Nyhan syndrome, schizophrenia, cerebellar
 CC ataxia, pain and alcoholism. They may also be used as immunogens to
 CC produce antibodies specific for the invention, and as vaccines.
 CC Transgenic cells containing a NOVX expressing construct are useful to
 CC produce non-human transgenic animals for studying the function and/or
 CC activity of the NOVX proteins and for identifying and/or evaluating
 CC modulators of NOVX protein activity. Transgenic cells containing a NOVX
 CC expressing construct are useful to produce non-human transgenic animals
 CC for studying the function and/or activity of the NOVX proteins and for
 CC identifying and/or evaluating modulators of NOVX protein activity. The
 CC sequences presented in ABS63431-ABS63444 are the human NOV1-NOV8 cDNAs
 XX

Sequence 6201 BP; 1736 A; 1552 C; 1707 G; 1206 T; 0 U; 0 Other;
 SQ

Query Match 99.0%; Score 6105.4; DB 6; Length 6201;
 Best Local Similarity 99.6%; Pred. No. 0;
 Matches 6131; Conservative 0; Mismatches 21; Indels 3; Gaps 1;

QY 1 ATGTTGAAGTTCATAATATGAGAGCGCGGATCTCTTGGAGTGTGCTGTGAACCAT 60
 DB 1 ATGTTGAAGTTCATAATATGAGAGCGCGGATCTCTTGGAGTGTGCTGTGAACCAT 60
 QY 61 GCCAGCGGCGCTCCAGGCTGATCTGTTCTTCCAGGGGAAACCACTTTATGACTCAA 120
 DB 61 GCCAGCGGCGCTCCAGGCTGATCTGTTCTTCCAGGGGAAACCACTTTATGACTCAA 120
 QY 121 CAGCAGATGTCCTCTTCCGAGAGGGATATTAGATGCCCTTTGTTCTTTGAA 180
 DB 121 CAGCAGATGTCCTCTTCCGAGAGGGATATTAGATGCCCTTTGTTCTTTGAA 180
 QY 181 GAATGAGTCAAGCAGCTGCTCTGATGAAGATTAGACAGTGAACATTTCCGGAAGTAT 240
 DB 181 GAATGAGTCAAGCAGCTGCTCTGATGAAGATTAGACAGTGAACATTTCCGGAAGTAT 240
 QY 241 TCCGACACCATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 300
 DB 241 TCCGACACCATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 300
 QY 301 AGCTTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 360
 DB 301 AGCTTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 360
 QY 361 GACATCTATGCTATGATAAAGATGATAAAGAGGCTTTATGAGCCAGAGAGAGGTTTCA 420
 DB 361 GACATCTATGCTATGATAAAGATGATAAAGAGGCTTTATGAGCCAGAGAGAGGTTTCA 420

Db 2578 CTGAGACACAGGCTGGAAAGTTGGAAGCCCAAGAACCGAATACTGGAGAGACAGCTGGAG 2637
Qy 2641 AAGATCACCAACCAAGACCAACAGTGAACAAGATGGCTGTGTGAATCTGAGACAAAGATTG 2700
Db 2638 AAGATCAGCCACCAAGACCAACAGTGAACAAGATGGCTGTGTGAATCTGAGACAAAGATTG 2697
Qy 2701 GGGAGGTCAGTCTGAGAGACGAGAGACAAAGAAATCTGAGAGCTCAAGCCCAAGTCAAGAG 2760
Db 2698 GGGAGGTCAGTCTGAGAGACGAGAGACAAAGAAATCTGAGAGCTCAAGCCCAAGTCAAGAG 2757
Qy 2761 CTAAGCTCTCCCTGCAAGAGCCGAGTCAAGTTGACAGCCCTGCAAGCTGCAAGAG 2820
Db 2758 CTAAGCTCTCCCTGCAAGAGCCGAGTCAAGTTGACAGCCCTGCAAGCTGCAAGAG 2817
Qy 2821 GCCCTGAGAGCCAGCTTCCGCAAGGCGAAGACAGAGCTGGAAGAGACCAAGACAAAGCT 2880
Db 2818 GCCCTGAGAGCCAGCTTCCGCAAGGCGAAGACAGAGCTGGAAGAGACCAAGACAAAGCT 2877
Qy 2881 GAAGAGAGATCCAGAGCTCAAGGCACTAAGAGATGAATAATCCAGCCCAATTTGATGCT 2940
Db 2878 GAAGAGAGATCCAGAGCTCAAGGCACTAAGAGATGAATAATCCAGCCCAATTTGATGCT 2937
Qy 2941 CTTGATTAACAGCTGATCTGTATCTACAGACCTGGAAGAGAGCTAAACAGCTGACCGAG 3000
Db 2938 CTTGATTAACAGCTGATCTGTATCTACAGACCTGGAAGAGAGCTAAACAGCTGACCGAG 2997
Qy 3001 GACAAAGCTGAATCAACAAACAACTTTGATCTGATCAACAACTGATGAGCTTCT 3060
Db 2998 GACAAAGCTGAATCAACAAACAACTTTGATCTGATCAACAACTGATGAGCTTCT 3057
Qy 3061 GGGGCCAACGACGAGATTGTACAACTGCGAAGTGAAGTGAACATCTCCCGGAGATC 3120
Db 3058 GGGGCCAACGACGAGATTGTACAACTGCGAAGTGAAGTGAACATCTCCCGGAGATC 3117
Qy 3121 ACGGAACGAGAGATGACAGCTTACAGCCAGAGCAAAACGATGAGGCTGGAAGACCAAG 3180
Db 3118 ACGGAACGAGAGATGACAGCTTACAGCCAGAGCAAAACGATGAGGCTGGAAGACCAAG 3177
Qy 3181 TGCACCATCTGAGAGAAACAGTCAATGATTTGAGAGCCCTAAACGATGAGCTGAGAA 3240
Db 3178 TGCACCATCTGAGAGAAACAGTCAATGATTTGAGAGCCCTAAACGATGAGCTGAGAA 3237
Qy 3241 AAAAGACGGCAGTGGAGAGCCCTGGAAGAGCGTCTTGAGTGAAGAAATCCAGTTGAG 3300
Db 3238 AAAAGACGGCAGTGGAGAGCCCTGGAAGAGCGTCTTGAGTGAAGAAATCCAGTTGAG 3297
Qy 3301 TGTGCGGTTGAGAGCTGCAAGAGATGCTGGAACACGAGAAACAGACAGGCGGAGAGCC 3360
Db 3298 TGTGCGGTTGAGAGCTGCAAGAGATGCTGGAACACGAGAAACAGACAGGCGGAGAGCC 3357
Qy 3361 GATCAGCGGATCACCGAGTCTGCGCAAGTGTGAGAGCTGCAAGTGAAGACCAAGGCT 3420
Db 3358 GATCAGCGGATCACCGAGTCTGCGCAAGTGTGAGAGCTGCAAGTGAAGACCAAGGCT 3417
Qy 3421 GAGATTCTGCTGTGAGAGGCTCTCAAGAGACAGAGCTGAAGGCCCAAGACCTTCT 3480
Db 3418 GAGATTCTGCTGTGAGAGGCTCTCAAGAGACAGAGCTGAAGGCCCAAGACCTTCT 3477
Qy 3481 GACAAGCTCAATGACCTGAGAAAGAGCAGTGCATGCTTGAATGAATGCCAGAGCTTA 3540
Db 3478 GACAAGCTCAATGACCTGAGAAAGAGCAGTGCATGCTTGAATGAATGCCAGAGCTTA 3537
Qy 3541 CAGCAGAAAGCTGAGAGCTGAACGAGAGCTCAACAGAGGCTTCTGGAAGAGCAAGCAAA 3600
Db 3538 CAGCAGAAAGCTGAGAGCTGAACGAGAGCTCAACAGAGGCTTCTGGAAGAGCAAGCAAA 3597
Qy 3601 TTAACAGACAGATGAGCCTGCAAGAAATCAATTTTCGTTGACTCAAGAGCTGCAA 3660
Db 3598 TTAACAGACAGATGAGCCTGCAAGAAATCAATTTTCGTTGACTCAAGAGCTGCAA 3657
Qy 3661 GAAAGCTTAAGATCGGCTGATCTTACTGAAGACAAAGAAAGTGAATGAGATCAAGCTG 3720

Db 3658 GAAAGCTTAAGATCGGCTGATCTTACTGAAGACAAAGAAAGTGAATGAGATATCAGCTG 3717
Qy 3721 GAAAAACATTCAGGTTCTATATTCATGAAAGAGTGAAGAAATGAAAGCACTATTTCTCAA 3780
Db 3718 GAAAAACATTCAGGTTCTATATTCATGAAAGAGTGAAGAAATGAAAGCACTATTTCTCAA 3777
Qy 3781 CAAACCAATCTCATTTGATTTTCTGCAAGCCCAAAATGAGACCACTGCTTAAAGAAAAAG 3840
Db 3778 CAAACCAATCTCATTTGATTTTCTGCAAGCCCAAAATGAGACCACTGCTTAAAGAAAAAG 3837
Qy 3841 GTTCTCTGAGATACAAATGAGAGTGAAGCTGAGCCCTGGAAGAGAAAGCTGCTGAGCA 3900
Db 3838 GTTCTCTGAGATACAAATGAGAGTGAAGCTGAGCCCTGGAAGAGAAAGCTGCTGAGCA 3897
Qy 3901 GAGCTAGAGAGACCTTCTGAGAAAGCCGATCGAGCTCGGTTCCGCGGAGAGAGCT 3960
Db 3898 GAGCTAGAGAGACCTTCTGAGAAAGCCGATCGAGCTCGGTTCCGCGGAGAGAGCT 3957
Qy 3961 GCCCACCGCAAGCAAGGACCAACCCATCCATCCAGCCAGCCACCGGAGAGCAG 4020
Db 3958 GCCCACCGCAAGCAAGGACCAACCCATCCAGCCAGCCACCGGAGAGCAG 4017
Qy 4021 ATGCGCATGTCGCGCATTCGTCGCGGTCGAGAGACCAAGCCAGTGCATGAGCGTCTG 4080
Db 4018 ATGCGCATGTCGCGCATTCGTCGCGGTCGAGAGACCAAGCCAGTGCATGAGCGTCTG 4077
Qy 4081 GCCCGCCCATCCAGCCGAGAGAAAGAGTCTTCAACTCCAGAGAGAAATTAATCCGCTCTT 4140
Db 4078 GCCCGCCCATCCAGCCGAGAGAAAGAGTCTTCAACTCCAGAGAGAAATTAATCCGCTCTT 4137
Qy 4141 AAGGAACGATGACCAACCAATATTCCTCACCGATTCAAGTGAAGTGAACATGCAAGCC 4200
Db 4138 AAGGAACGATGACCAACCAATATTCCTCACCGATTCAAGTGAAGTGAACATGCAAGCC 4197
Qy 4201 ACAAGATGTCGTGTGTGTCGGAATACGTCACCTTTGGAAGCGCAGCATCAATGTCCTC 4260
Db 4198 ACAAGATGTCGTGTGTGTCGGAATACGTCACCTTTGGAAGCGCAGCATCAATGTCCTC 4257
Qy 4261 GAATGTCAGGTATGTGTGTCACCCCAAGTCTCCACGTCGTTGCAAGCCACTGCGCTTG 4320
Db 4258 GAATGTCAGGTATGTGTGTCACCCCAAGTCTCCACGTCGTTGCAAGCCACTGCGCTTG 4317
Qy 4321 CTTGCTGAATATGTCACACACATTCACCGAGGCTTCTGCGCTGACAAATGAACTCCCA 4380
Db 4318 CTTGCTGAATATGTCACACACATTCACCGAGGCTTCTGCGCTGACAAATGAACTCCCA 4377
Qy 4381 GGTCTCCAGACCAAGAGGCTCCAGAGCACTTGCACCTGGAAGGGGTGAAGAGTGGCC 4440
Db 4378 GGTCTCCAGACCAAGAGGCTCCAGAGCACTTGCACCTGGAAGGGGTGAAGAGTGGCC 4437
Qy 4441 AGGAATTAACAAACGAGAGCAGAGAGGCTGGGACAGGAATGACTTGTCTGAGAGGATCA 4500
Db 4438 AGGAATTAACAAACGAGAGCAGAGAGGCTGGGACAGGAATGACTTGTCTGAGAGGATCA 4497
Qy 4501 AAAATGCTCATTTTATGACATGTAAGCCAGAGAGCTGGAACAGAGCCGCTGGAAGATT 4560
Db 4498 AAAATGCTCATTTTATGACATGTAAGCCAGAGAGCTGGAACAGAGCCGCTGGAAGATT 4557
Qy 4561 GAGCTGTGCTTCCGACGGGAGATGATATTAATCATGTGCTGCTGTGCTTCCGAATC 4620
Db 4558 GAGCTGTGCTTCCGACGGGAGATGATATTAATCATGTGCTGCTGTGCTTCCGAATC 4617
Qy 4621 GCAAAATACAGCCAAAGAGATGTCATATCTGAAGATGGAATCTCACCCGCAAC 4680
Db 4618 GCAAAATACAGCCAAAGAGATGTCATATCTGAAGATGGAATCTCACCCGCAAC 4677
Qy 4681 ACTGCTGGCCCGGAGAAACCTCTACTGCTAGCTCCCAAGCTTCCCTGCAAAACAGGCG 4740
Db 4678 ACTGCTGGCCCGGAGAAACCTCTACTGCTAGCTCCCAAGCTTCCCTGCAAAACAGGCG 4737
Qy 4741 TGGGTCAACGCTTGAATCAATGTGTCAGAGTGGAGAGTTTCTAAGGAAAAACAGAA 4800
Db 4738 TGGGTCAACGCTTGAATCAATGTGTCAGAGTGGAGAGTTTCTAAGGAAAAACAGAA 4797

QY 4801 GCTGATGCTAACTGCTTGGAACCTCCCTGCTGAACTGGAAGGTGATGACCCGTCTAGAC 4860
 Db 4798 GCTGATGCTAACTGCTTGGAACCTCCCTGCTGAACTGGAAGGTGATGACCCGTCTAGAC 4857
 QY 4861 ATGAACTGCAACGCTGCTGCTTCACTGACACGAGTGTGTTGGTGGGACACCGAGAAAGGGCTC 4920
 Db 4858 ATGAACTGCAACGCTGCTGCTTCACTGACACGAGTGTGTTGGTGGGACACCGAGAAAGGGCTC 4917
 QY 4921 TAGCCCTGAATGCTTGAAGAACTCCCTAACCATATGCTCCAGGAATTTGAGAGAGTCTTC 4980
 Db 4918 TAGCCCTGAATGCTTGAAGAACTCCCTAACCATATGCTCCAGGAATTTGAGAGAGTCTTC 4977
 QY 4981 CAAATTTATATTAATCAAGAACCTGAGAGACTACTCATGATAGCAGAGAAAGACGGGCA 5040
 Db 4978 CAAATTTATATTAATCAAGAACCTGAGAGACTACTCATGATAGCAGAGAAAGACGGGCA 5037
 QY 5041 CTGTGCTTTGTGAGAGCTGGAAGAAAGTAAACAGTCCCTGGCCAGTCCCACTGCTGCC 5100
 Db 5038 CTGTGCTTTGTGAGAGCTGGAAGAAAGTAAACAGTCCCTGGCCAGTCCCACTGCTGCC 5097
 QY 5101 CAGCCGACACTCACCACCAATTTTGAAGCTGTCAAGGGGCTGCCATTGTTGGGGCA 5160
 Db 5098 CAGCCGACACTCACCACCAATTTTGAAGCTGTCAAGGGGCTGCCATTGTTGGGGCA 5157
 QY 5161 GGGCAAGTTGAGAACGGGCTCTGCATCTGACAGCCATGCCCAAGAGTGTCAATTC 5220
 Db 5158 GGGCAAGTTGAGAACGGGCTCTGCATCTGACAGCCATGCCCAAGAGTGTCAATTC 5217
 QY 5221 CGCTACACAGAAACCTCAGCAAAATCTGCATCCGGAAGAGATAGAGACCTCAGAGCCC 5280
 Db 5218 CGCTACACAGAAACCTCAGCAAAATCTGCATCCGGAAGAGATAGAGACCTCAGAGCCC 5277
 QY 5281 TGAAGCTGTATCACTTCAACATTAAGATCCTCATTTGGAACCAATAATTTAGAA 5340
 Db 5278 TGAAGCTGTATCACTTCAACATTAAGATCCTCATTTGGAACCAATAATTTAGAA 5337
 QY 5341 ATGACACTGAGAGAGTACAGCTCAGAGAAATCTGTGATTAAGATGACATTCCTTGGCA 5400
 Db 5338 ATGACACTGAGAGAGTACAGCTCAGAGAAATCTGTGATTAAGATGACATTCCTTGGCA 5397
 QY 5401 CCTGCTGTGTTGGCCGCTCTTCCACAGCTTCCCTGTCTCAATGCTGACAGTGAACAGC 5460
 Db 5398 CCTGCTGTGTTGGCCGCTCTTCCACAGCTTCCCTGTCTCAATGCTGACAGTGAACAGC 5457
 QY 5461 GCAGGGCAGAGAGAGAGTACTGCTGTGTTCCACGAATTTGAGAGTGTCTGTGATTC 5520
 Db 5458 GCAGGGCAGAGAGAGAGTACTGCTGTGTTCCACGAATTTGAGAGTGTCTGTGATTC 5517
 QY 5521 TACGGAAGAGTACGCGCAGACAGATCTCAAGTGAAGTGCCTTACCTTTGGCCTTTGCC 5580
 Db 5518 TACGGAAGAGTACGCGCAGACAGATCTCAAGTGAAGTGCCTTACCTTTGGCCTTTGCC 5577
 QY 5581 TACAGAGAACCTATCTGTTTGTGACCCACTTCAACTCTCGAAGTAATTTAGATCCAG 5640
 Db 5578 TACAGAGAACCTATCTGTTTGTGACCCACTTCAACTCTCGAAGTAATTTAGATCCAG 5637
 QY 5641 GCACGCTCTCAGAGAGGAGACCCCTGCCGAGCGTACTGTGACATCCCGAACCCGCGTAC 5700
 Db 5638 GCACGCTCTCAGAGAGGAGACCCCTGCCGAGCGTACTGTGACATCCCGAACCCGCGTAC 5697
 QY 5701 CTGGGACCTGCAATTTCTCTCAGAGAGATTTACTTGGCGTCTCATACAGATTAATTA 5760
 Db 5698 CTGGGACCTGCAATTTCTCTCAGAGAGATTTACTTGGCGTCTCATACAGATTAATTA 5757
 QY 5761 AGGATCATTTGCTGCAAGGAGAACTCTGTGAAGAGTCCGCACTGAACCAACCGGGGC 5820
 Db 5758 AGGATCATTTGCTGCAAGGAGAACTCTGTGAAGAGTCCGCACTGAACCAACCGGGGC 5817
 QY 5821 CCGTCCACCTTCCGACAGAGCCCAACAGAGAGCCCAACAGTACAAAGAGACATC 5880
 Db 5818 CCGTCCACCTTCCGACAGAGCCCAACAGAGAGCCCAACAGTACAAAGAGACATC 5877

QY 5881 ACCAAGCGCTGCTCTTCAAGCCCAAGCGCCGCCCGGAAGGCCCAAGCCACCCCGAGAGCCA 5940
 Db 5878 ACCAAGCGCTGCTCTTCAAGCCCAAGCGCCGCCCGGAAGGCCCAAGCCACCCCGAGAGCCA 5937
 QY 5941 AGGACACCCCAACCGCTACCGCGAGGGGCGGACCGAGCTGCGCAGAGGCAAGTCTCTGGC 6000
 Db 5938 AGGACACCCCAACCGCTACCGCGAGGGGCGGACCGAGCTGCGCAGAGGCAAGTCTCTGGC 5997
 QY 6001 CGCCCTGAGAGAGAGAAATCCCGCGCGGATATCTAGCAACGCGAGAGAGCGGTCC 6060
 Db 5998 CGCCCTGAGAGAGAGAAATCCCGCGCGGATATCTAGCAACGCGAGAGAGCGGTCC 6057
 QY 6061 CCGCGAGGCTGTTTGAAGACAGCAGAGGGGCGGCTGCTGCGGAGCCGTGAAGACC 6120
 Db 6058 CCGCGAGGCTGTTTGAAGACAGCAGAGGGGCGGCTGCTGCGGAGCCGTGAAGACC 6117
 QY 6121 CCGCTGCCAGGTGAACAGGTCTGGGACAGTC 6155
 Db 6118 CCGCTGCCAGGTGAACAGGTGAGGACAGCATTC 6152

RESULT 12
 ADA05653
 ID ADA05653 standard; cDNA; 6201 BP.
 AC ADA05653;
 XX
 DT 06-NOV-2003 (first entry)
 DE Human NOV1g encoding cDNA SEQ ID NO:13.
 XX
 KW human; NOV1; antidiabetic; anorectic; antibacterial; virucide;
 KW immunomodulator; cytostatic; nootropic; neuroprotective;
 KW antiparkinsonian; antilipemic; gene therapy; human disease;
 KW metabolic disorder; diabetes; obesity; infection; cachexia; cancer;
 KW neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;
 KW immune disorder; haematopoietic disorder; dyslipidaemia; gene; ss.
 XX
 OS Homo sapiens.
 FH Key. Location/Qualifiers
 FT CDS 1..6201
 FT /*tag= a
 FT /product= "NOV1g"
 XX
 PN M02003029424-A2.
 PD 10-APR-2003.
 XX
 PF 02-OCT-2002; 2002MO-US031373.
 XX
 PR 02-OCT-2001; 2001US-0326435P.
 PR 05-OCT-2001; 2001US-0327435P.
 PR 05-OCT-2001; 2001US-0327449P.
 PR 09-OCT-2001; 2001US-0327917P.
 PR 09-OCT-2001; 2001US-0328029P.
 PR 09-OCT-2001; 2001US-0328044P.
 PR 09-OCT-2001; 2001US-0328056P.
 PR 12-OCT-2001; 2001US-0328849P.
 PR 15-OCT-2001; 2001US-0329414P.
 PR 17-OCT-2001; 2001US-0330142P.
 PR 18-OCT-2001; 2001US-0330309P.
 PR 22-OCT-2001; 2001US-0341058P.
 PR 24-OCT-2001; 2001US-0339266P.
 PR 24-OCT-2001; 2001US-0343629P.
 PR 29-OCT-2001; 2001US-0349575P.
 PR 01-NOV-2001; 2001US-0346357P.
 PR 17-APR-2002; 2002US-0373260P.
 PR 19-APR-2002; 2002US-0373815P.
 PR 19-APR-2002; 2002US-0373817P.
 PR 19-APR-2002; 2002US-0373826P.
 PR 19-APR-2002; 2002US-0373884P.
 PR 22-APR-2002; 2002US-0374977P.

PR 16-MAY-2002; 2002US-0381037P.
PR 16-MAY-2002; 2002US-0381038P.
PR 16-MAY-2002; 2002US-0381042P.
PR 17-MAY-2002; 2002US-0381642P.
PR 28-MAY-2002; 2002US-0383656P.
PR 29-MAY-2002; 2002US-0383831P.
PR 25-JUN-2002; 2002US-0391335P.
PR 01-OCT-2002; 2002US-00262511.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Smithson G, Millet I, Peyman JA, Kekuda R, Ju J, Li L, Guo X;
PI Paturajan M, Spytek KA, Edinger SR, Ellemann K, Malysankar UM;
PI Ott T, Gorman L, Zethusen BD, Anderson DW, Zhong M, Catterton E;
PI Ji W, Miller CE, Raestelli L, Stone DJ, Pena CBA, Shenoy SG;
PI Shimets RA, Rothenberg ME, Leach MD, Agee MU, Berghs C, Dipippo VA;
PI Eisen AJ, Gangoli EA, Rieger DK, Spaderna SK;
XX
DR WPI; 2003-381626/36.
DR P-PSDB; ADA05654.
XX
PT New NOXV polypeptides and nucleic acids, useful for diagnosing,
PT preventing or treating NOXV-associated disorders, e.g. diabetes, obesity,
PT cancer or dyslipidemia, and in chromosome mapping, tissue typing or
PT pharmacogenomics.
XX
PS Claim 20, Page 104-105, 586pp; English.
XX
CC The present invention describes NOXV proteins, where X can be 1 to 55
CC (e.g. NOV1). Also described: (1) a composition comprising a polypeptide
CC described above and a carrier; (2) a kit comprising, in one or more
CC containers, the composition described above; (3) an isolated nucleic acid
CC molecule which encodes a NOXV protein of the invention; (4) a vector
CC comprising the nucleic acid molecule described above; (5) a cell
CC comprising the above vector; (6) an antibody that immunospecifically
CC binds to the polypeptide described above; (7) methods for determining the
CC presence or amount of the above polypeptide or nucleic acid molecule in a
CC sample; (8) methods for determining the presence of or predisposition to
CC a disease associated with altered levels of expression of the above
CC polypeptide or nucleic acid molecule in a first mammalian subject; (9) a
CC method of identifying an agent that binds to the polypeptide described
CC above; (10) a method for identifying a potential therapeutic agent for
CC use in treating a pathology that is related to an aberrant expression or
CC aberrant physiological interactions of the polypeptide; (11) a method of
CC screening for a modulator of activity or of latency or predisposition to
CC a pathology associated with the polypeptide; (12) a method for modulating
CC the activity of the polypeptide described above; (13) methods for treating
CC or preventing a pathology associated with the above polypeptide in a
CC mammal; and (14) a method for producing the above polypeptide, NOXV
CC sequences have antidiabetic, anorectic, antibacterial, vitruce, and
CC immunomodulator, cytostatic, nootropic, neuroprotective, antiparkinsonian
CC and anti-lipemic activities, and can be used in gene therapy. The
CC polypeptide is useful in manufacturing a medicament for treating a
CC syndrome associated with a human disease. The polypeptide or the nucleic
CC acid molecule may be used to diagnose, treat or prevent metabolic
CC disorders such as diabetes or obesity, infections, cachexia, cancer,
CC neurodegenerative disorders such as Alzheimer's disease or Parkinson's
CC disease, immune disorders, haematopoietic disorders and various
CC dyslipidaemias. The nucleic acids can also be used as hybridisation
CC probes, in chromosome mapping, tissue typing, preventive medicine and
CC pharmacogenomics. The present sequence encodes a human NOXV protein from
CC the present invention.
XX
SQ Sequence 6201 BP; 1736 A; 1552 C; 1707 G; 1206 T; 0 U; 0 Other;

Query Match 99.0%; Score 6105.4; DB 8; Length 6201;
Best Local Similarity 99.6%; Pred No. 0;
Matches 6131; Conservative 0; Mismatches 21; Indels 3; Gaps 1;

Qy 1 ATGTTGAAGTCAAAATATGAGCGCGGAATCCTTTGGATGCTGCTGCTGAACCAT 60
Db 1 ATGTTGAAGTCAAAATATGAGCGCGGAATCCTTTGGATGCTGCTGCTGAACCAT 60

Qy 61 GCCAGCCGGGCTTCAGGCTGAATCTGTCTTCAGGGGAAACCAACCTTTATGACTCA 120
Db 61 GCCAGCCGGGCTTCAGGCTGAATCTGTCTTCAGGGGAAACCAACCTTTATGACTCA 120
Qy 121 CAGCAGATGCTCTCTCTTTCCCGAAGAGGATATAGATGCCCTTGTCTTGTGAA 180
Db 121 CAGCAGATGCTCTCTCTTTCCCGAAGAGGATATAGATGCCCTTGTCTTGTGAA 180
Qy 181 GAATGACGAGCTGCTCTGATGAAGATTAGCAAGTGAACAATTGTCCGGAATAT 240
Db 181 GAATGACGAGCTGCTCTGATGAAGATTAGCAAGTGAACAATTGTCCGGAATAT 240
Qy 241 TCCGACACCATTAAGTATGAGAGCTCCAGCTTCGCAAGAGCTTGAAGTACA 300
Db 241 TCCGACACCATTAAGTATGAGAGCTCCAGCTTCGCAAGAGCTTGAAGTACA 300
Qy 301 AGCTTGTAGTGTGTGCTCACTTTGCTGAAGTCAAGTGTGTGAAGAGAAACCGGG 360
Db 301 AGCTTGTAGTGTGTGCTCACTTTGCTGAAGTCAAGTGTGTGAAGAGAAACCGGG 360
Qy 361 GACATCTATGCTATGAAAGTATGAAGAAAGGCTTTATTTGGCCAGAGAGATTCA 420
Db 361 GACATCTATGCTATGAAAGTATGAAGAAAGGCTTTATTTGGCCAGAGAGATTCA 420
Qy 421 TTTTGGAGAAAGCGGAAACATTTATCTCGAAGCAACAGCCCTGTGATCCCCAA 480
Db 421 TTTTGGAGAAAGCGGAAACATTTATCTCGAAGCAACAGCCCTGTGATCCCCAA 480
Qy 481 CAGTATGCTTTTCAAGCAAAATCACTTTATCTGTGATGAGATGAGCTTGAAGG 540
Db 481 CAGTATGCTTTTCAAGCAAAATCACTTTATCTGTGATGAGATGAGCTTGAAGG 540
Qy 541 GACTTCTGTCACTTTTGAATGATGAGAACAGTATGATGAAACCTGATACAGTT 600
Db 541 GACTTCTGTCACTTTTGAATGATGAGAACAGTATGATGAAACCTGATACAGTT 600
Qy 601 TACCTGCTGAGTATTTTGGCTGTTCACAGGTTTCACTGATGAGAAACGTCAG 660
Db 601 TACCTGCTGAGTATTTTGGCTGTTCACAGGTTTCACTGATGAGAAACGTCAG 660
Qy 661 GACATCAAGCTGAGAAACATTCCTGTTGACCGACAGCAATCAAGCTGTGATTT 720
Db 661 GACATCAAGCTGAGAAACATTCCTGTTGACCGACAGCAATCAAGCTGTGATTT 720
Qy 721 GGAATGCGCGGAAATGAATCAACAGATGTGAATGCCAACTCCGATTGGAGC 780
Db 721 GGAATGCGCGGAAATGAATCAACAGATGTGAATGCCAACTCCGATTGGAGC 780
Qy 781 CCAGATTATAGCTGCTTGAAGTGTGATGAGTGAAGAGGAGTGAAGGACCTTAC 840
Db 781 CCAGATTATAGCTGCTTGAAGTGTGATGAGTGAAGAGGAGTGAAGGACCTTAC 840
Qy 841 GGCCTGAGTGAAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 900
Db 841 GGCCTGAGTGAAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 900
Qy 897 GGCCTGAGTGAAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 960
Db 897 GGCCTGAGTGAAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 960
Qy 901 TCCCTTTCGAGAGGAACTCTGCGAAGACCTTCAATTAACATTATGATTTCCAG 960
Db 901 TCCCTTTCGAGAGGAACTCTGCGAAGACCTTCAATTAACATTATGATTTCCAG 960
Qy 957 TCCCTTTCGAGAGGAACTCTGCGAAGACCTTCAATTAACATTATGATTTCCAG 957
Db 957 TCCCTTTCGAGAGGAACTCTGCGAAGACCTTCAATTAACATTATGATTTCCAG 957
Qy 961 TTTTGAATTTTCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 1020
Db 961 TTTTGAATTTTCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 1020
Qy 958 TTTTGAATTTTCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 1017
Db 958 TTTTGAATTTTCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 1017
Qy 1021 TTTTGAATTTTCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 1080
Db 1021 TTTTGAATTTTCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 1080
Qy 1018 TTTTGAATTTTCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 1077
Db 1018 TTTTGAATTTTCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 1077
Qy 1081 TTTTGAATTTTCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 1140
Db 1081 TTTTGAATTTTCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 1140
Qy 1078 TTTTGAATTTTCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 1137
Db 1078 TTTTGAATTTTCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 1137
Qy 1141 TTTTGAATTTTCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 1200
Db 1141 TTTTGAATTTTCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 1200

```
Db 1138 TCTGACATATACACCTCAATTTTGATGAACAGAGAAAGAAATTCGTGGGTTTCACTCTC,1197
Qy 1201 CCGTGCCAGCTGAGCCCCCTCAGGCTTCTCGGATGAAGAACTGCCGTTTGAGGGTTTTCG 1260
Db 1198 CCGTGCCAGCTGAGCCCCCTCAGGCTTCTCGGATGAAGAACTGCCGTTTGAGGGTTTTCG 1257
Qy 1261 TACAGCAAGGCACTGGGGATTTCTTGATGAATCTGAGTCTGTTGTGTGTCGGGTCTGGACTCC 1320
Db 1258 TACAGCAAGGCACTGGGGATTTCTTGATGAATCTGAGTCTGTTGTGTGTCGGGTCTGGACTCC 1317
Qy 1321 CTTGCCAAGACTAGCTCCATGAGAAAAGAACTTCTATCAAAAGCAAAAGACTCAAGAC 1380
Db 1318 CTTGCCAAGACTAGCTCCATGAGAAAAGAACTTCTATCAAAAGCAAAAGACTCAAGAC 1377
Qy 1381 TCTCAGAGCAAGTGTCAAGAGATGAGAGAAATGACCCGGTTACATCCGAGAGTGTCA 1440
Db 1378 TCTCAGAGCAAGTGTCAAGAGATGAGAGAAATGACCCGGTTACATCCGAGAGTGTCA 1437
Qy 1441 GAGGTGAGGCTGTGCTTAGTCAAGAGAGAGGTGAGCTGAAGGCTCTGAGACTCAGAGA 1500
Db 1438 GAGGTGAGGCTGTGCTTAGTCAAGAGAGAGGTGAGGCTGAAGGCTCTGAGACTCAGAGA 1497
Qy 1501 TCCCTCTGAGAGCAAGGACTTTGCTAATCAATCAAGATGCAATGCAATGAGCTTAAGCGAAGT 1560
Db 1498 TCCCTCTGAGAGCAAGGACTTTGCTAATCAATCAAGATGCAATGCAATGAGCTTAAGCGAAGT 1557
Qy 1561 TTGGAGCAAGCAAGAGTGAAGGTGTCCAGAGAGATGCAAGCACTGCAAGCTTCTCAT 1620
Db 1558 TTGGAGCAAGCAAGAGTGAAGGTGTCCAGAGAGATGCAAGCACTGCAAGCTTCTCAT 1617
Qy 1621 GATATCAGAGAGCAAGAGCCGGAAGCTCAAGAAATCAAGAGCAAGAGTACAGGCTCAA 1680
Db 1618 GATATCAGAGAGCAAGAGCCGGAAGCTCAAGAAATCAAGAGCAAGAGTACAGGCTCAA 1677
Qy 1681 GTGGAGAAATGAGGTTGATGATGAATCAAGTTGGAAGAGATCTTGTCTCAGCAAGAA 1740
Db 1678 GTGGAGAAATGAGGTTGATGATGAATCAAGTTGGAAGAGATCTTGTCTCAGCAAGAA 1737
Qy 1741 CGGAGTATCTCTACGAATCTGAGCTGAGAGAGTCTGGCTGTGCTGAAGAAATTCAG,1800
Db 1738 CGGAGTATCTCTACGAATCTGAGCTGAGAGAGTCTGGCTGTGCTGAAGAAATTCAG,1797
Qy 1801 CGGAAGCGACAGAGATGTCAAGATAAATCTGTTGAAGCTTAAGATCAAGGGAAGCTGAA 1860
Db 1798 CGGAAGCGACAGAGATGTCAAGATAAATCTGTTGAAGCTTAAGATCAAGGGAAGCTGAA 1857
Qy 1861 GTGGAGAAATATGCGAAATGAGAGAGATCAATGCTGAGCAGCAGCTCAAAATTCAGAG 1920
Db 1858 GTGGAGAAATATGCGAAATGAGAGAGATCAATGCTGAGCAGCAGCTCAAAATTCAGAG 1917
Qy 1921 CTTCAAGAGAAATGAGAGAGGCTGTAAAGCCAGAGCAGAGGCTACCGAGCTGCTCAG 1980
Db 1918 CTTCAAGAGAAATGAGAGAGGCTGTAAAGCCAGAGCAGAGGCTACCGAGCTGCTCAG 1977
Qy 1981 AATATCCGAGGCAAGAGAGAGCGAGCGAGGGAGCTGAGAAAGCTGCGAAGCCAGAG 2040
Db 1978 AATATCCGAGGCAAGAGAGAGCGAGCGAGGGAGCTGAGAAAGCTGCGAAGCCAGAG 2037
Qy 2041 GATTCTTCTGAAGGCAATCAGAAAGAGCTGTGTGAGCTGAGAAACGCCCATTTCTCTG 2100
Db 2038 GATTCTTCTGAAGGCAATCAGAAAGAGCTGTGTGAGCTGAGAAACGCCCATTTCTCTG 2097
Qy 2101 GAGAGCAAGTAAAGAGACTTGAAGACCATGAGGCTGTGAAGAAAACAATCGAAGAGATGAC 2160
Db 2098 GAGAGCAAGTAAAGAGACTTGAAGACCATGAGGCTGTGAAGAAAACAATCGAAGAGATGAC 2157
Qy 2161 ATCCAGACAAATATCCCAAGATCCAGCAAGATGAGTGTGATATAAATTTCTGAGCTGAAGAG 2220
Db 2158 ATCCAGACAAATATCCCAAGATCCCAAGATGAGTGTGATATAAATTTCTGAGCTGAAGAG 2217
Qy 2221 AAACATGGAGGCGCAAGTCTCAGCCACAGCACTTAAGTGCACCTGAACAGAAAAG 2280
|||||
Db 2218 AAACATGGAGGCGCAAGTCTCAGCCACAGCACTTAAGTGCACCTGAACAGAAAAG 2277
Qy 2281 CAGCACTATGAGGAGAAAAGATTAAAGTTGAGCAATCAGATTAAGAAAGACTGGCTGAC 2340
Db 2278 CAGCACTATGAGGAGAAAAGATTAAAGTATGAGCAATCAGATTAAGAAAGACTGGCTGAC 2337
Qy 2341 AAGAGACACTGAGAGCAATGATGAGAGACACAGAGAGAGGCGCCATGAGAGGCGAAA 2400
Db 2338 AAGAGACACTGAGAGCAATGATGAGAGACACAGAGAGAGGCGCCATGAGAGGCGAAA 2397
Qy 2401 ATTCTCAGCAGACAGAAAGGCGATGATCAATGTGATTCGAAGATCCAGATCCCTGGAA 2460
Db 2398 ATTCTCAGCAGACAGAAAGGCGATGATCAATGTGATTCGAAGATCCCTGGAA 2457
Qy 2461 CAGAGATTTGGAATCTGTCTGAAGCCAAATTAATCTGACAAATAGACATCTTTTACC 2520
Db 2458 CAGAGATTTGGAATCTGTCTGAAGCCAAATTAATCTGACAAATAGACATCTTTTACC 2517
Qy 2521 CAAAGGAACTGAAGGCGCAAGAGAGATGATTTCTGAATCTCAGGCAACAGAAATTTTAC 2580
Db 2518 CAAAGGAACTGAAGGCGCAAGAGAGATGATTTCTGAATCTCAGGCAACAGAAATTTTAC 2577
Qy 2581 CTGAGACACAGGCTGGGAAGTTGAGAGCCAGAACCGAAATCTGAGAGGACGCTGGAG 2640
Db 2578 CTGAGACACAGGCTGGGAAGTTGAGAGCCAGAACCGAAATCTGAGAGGACGCTGGAG 2637
Qy 2641 AAGATCAGCCACCAAGACCAAGATGACCAAGATTCGCTGTGGAATCTGAGACAAAGATTG 2700
Db 2638 AAGATCAGCCACCAAGACCAAGATGACCAAGATTCGCTGTGGAATCTGAGACAAAGATTG 2697
Qy 2701 CGGAGGTCAGTCTTAGACACAGAGAGCAGAAACTGAGGCTCAAGCGCCAGCTTCAAGAG 2760
Db 2698 CGGAGGTCAGTCTTAGACACAGAGAGCAGAAACTGAGGCTCAAGCGCCAGCTTCAAGAG 2757
Qy 2761 CTACAGCTTCCCTGACAGAGCCCGAGTCAAGTTGACACCCCTGACAGGCTGACCGGCG 2820
Db 2758 CTACAGCTTCCCTGACAGAGCCCGAGTCAAGTTGACACCCCTGACAGGCTGACCGGCG 2817
Qy 2821 GCCCTGAGAGCCAGCTTGCACAGAGCGAAGACAGAGCTGGAAGAGACCAAGCAGAAAGCT 2880
Db 2818 GCCCTGAGAGCCAGCTTGCACAGAGCGAAGACAGAGCTGGAAGAGACCAAGCAGAAAGCT 2877
Qy 2881 GAAAGAGAGATCCAGGCACTCAGCGCACATAGAGATGAATCCAGCGCAAAATTTGATGCT 2940
Db 2878 GAAAGAGAGATCCAGGCACTCAGCGCACATAGAGATGAATCCAGCGCAAAATTTGATGCT 2937
Qy 2941 CTTGCTAACGCTGTACTGTATATCAAGACTTGAAGAGACGCTAAACCAAGCTGACCGAG 3000
Db 2938 CTTGCTAACGCTGTACTGTATATCAAGACTTGAAGAGACGCTAAACCAAGCTGACCGAG 2997
Qy 3001 GACAGGCTGAATCAACCAAGCAAAATCTTCACTGTCCAAACCAACTGAGTGAAGGCTTCT 3060
Db 2998 GACAGGCTGAATCAACCAAGCAAAATCTTCACTGTCCAAACCAACTGAGTGAAGGCTTCT 3057
Qy 3061 GCGCGCAACAGCAGATTTGTACAACCTGCAAGTGAAGTGAACCATCTCCGCGGGAGATC 3120
Db 3058 GCGCGCAACAGCAGATTTGTACAACCTGCAAGTGAAGTGAACCATCTCCGCGGGAGATC 3117
Qy 3121 ACGGAACGAGAGATGCAAGCTTACCAAGCAGAGCAAAAGATGAGGCTTGAAGACCAAG 3180
Db 3118 ACGGAACGAGAGATGCAAGCTTACCAAGCAGAGCAAAAGATGAGGCTTGAAGACCAAG 3177
Qy 3181 TGCACCATGCTGAGAGAAAGAGTCAATGGAATTTGAGAGGCTTAAACGATGAGCTGTGAA 3240
Db 3178 TGCACCATGCTGAGAGAAAGAGTCAATGGAATTTGAGAGGCTTAAACGATGAGCTGTGAA 3237
Qy 3241 AAAGAGCGGCAATGAGAGGCTTGAAGAGGCTCTGTGATGATGAAGAAATCCAGTTTGAG 3300
Db 3238 AAAGAGCGGCAATGAGAGGCTTGAAGAGGCTCTGTGATGATGAAGAAATCCAGTTTGAG 3297
Qy 3301 TGTGCGGTTTGAAGGCTGACAGAGATGCTTGAACACCGAAGAAACAGAGCAGAGGCGAGGCC 3360
Db 3298 TGTGCGGTTTGAAGGCTGACAGAGATGCTTGAACACCGAAGAAACAGAGCAGAGGCGAGGCC 3357
```

OY	3361	GATTCAGCCGATATCAACCGAGCTTCCGCCAGGTGTGTGAGCTGGCAGTGAAGAGCACAAGCT	3420
Db	3358	GATCAGCCGATATCAACCGAGCTTCCGCCAGGTGTGTGAGCTGGCAGTGAAGAGCACAAGCT	3417
OY	3421	GAGATTCTGCTCTCGCAGAGGCTCTCAAAAGCAGAAAGCTGAAGGCCAGAGCCTCTCT	3480
Db	3418	GAGATTCTGCTCTCGCAGAGGCTCTCAAAAGCAGAAAGCTGAAGGCCAGAGCCTCTCT	3477
OY	3481	GACAAAGCTCAATGACCTTGAGAAAGCAGTGTATGCTTGAATGATGCCGAAGCTTGA	3540
Db	3478	GACAAAGCTCAATGACCTTGAGAAAGCAGTGTATGCTTGAATGATGCCGAAGCTTGA	3537
OY	3541	CAGCAGAAAGCTTGAGAGCTGAAGAGGCTCAAAACAGAGGCTTCTGGAAGACAGAGCCAA	3600
Db	3538	CAGCAGAAAGCTTGAGAGCTGAAGAGGCTCAAAACAGAGGCTTCTGGAAGAGCAGAGCCAA	3597
OY	3601	TTACAGCAGCAGATGGAAGCTGAGAAAAATCAGATTTTCCTGTGACTCAAGAGCTGCA	3660
Db	3598	TTACAGCAGCAGATGGAAGCTGAGAAAAATCAGATTTTCCTGTGACTCAAGAGCTGCA	3657
OY	3661	GAACTCTAGATCGGGCTGATTTACTGAAGACAGAAAGAGTGACTTGGAGTATCAGTG	3720
Db	3658	GAAAGCTCTAGAGTCGGGCTGATCTACTGAAGACAGAAAGAGTGACTTGGAGTATCAGTG	3717
OY	3721	GAAGAACATTCAGGTTCTGATTTCTCATGAAAAAGTGAAGAGGACATATTCTCA	3780
Db	3718	GAAGAACATTCAGGTTCTGATTTCTCATGAAAAAGTGAAGAGGACATATTCTCA	3777
OY	3781	CAAGCAAACTCATTTGATTTTCTGCAAGCCAAATGACCAACTGCTAAAAAGAAAAAG	3840
Db	3778	CAAGCAAACTCATTTGATTTTCTGCAAGCCAAATGACCAACTGCTAAAAAGAAAAAG	3837
OY	3841	GTTCCCTGCACTACAAATGAGCTGAAGCTGGCCCTGGAGAAAGAAAGCTTCCTGTGCA	3900
Db	3838	GTTCCCTGCACTACAAATGAGCTGAAGCTGGCCCTGGAGAAAGAAAGCTTCCTGTGCA	3897
OY	3901	GAGCTAGAGAGCCCTTCAGAAAGCCCGCATCGAGCTCCGGCTCGGCCGGGAGAAAGCT	3960
Db	3898	GAGCTAGAGAGAGCCCTTCAGAAAGCCCGCATCGAGCTCCGGCTCGGCCGGGAGAAAGCT	3957
OY	3961	GCCCAACCGCAAGCAACGAGCACCACCAACCCATTCACAGCCAGCCAGCCAGGCAAG	4020
Db	3958	GCCCAACCGCAAGCAACGAGCACCACCAACCCATTCACAGCCAGCCAGGCAAGCAG	4017
OY	4021	ATCGCAATGTCCGCATCTGTGGGTGCGCAGAGCAAGACCCCACTGCAAGAGCCTCTG	4080
Db	4018	ATCGCAATGTCCGCATCTGTGGGTGCGCAGAGCAAGACCCCACTGCAAGAGCCTCTG	4077
OY	4081	GCCCCGCGCATCAGCCCGCAGAAAGAGTCTTCAACTCCAGAGAAATTAGTCGGCGCTT	4140
Db	4078	GCCCCGCGCATCAGCCCGCAGAAAGAGTCTTCAACTCCAGAGAAATTAGTCGGCGCTT	4137
OY	4141	AAGGAGGCGATGCAACCAATATTTCTCAACGATTCAACGTAAGACTGAATCGAGGCC	4200
Db	4138	AAGGAGGCGATGCAACCAATATTTCTCAACGATTCAACGTAAGACTGAATCGAGGCC	4197
OY	4201	ACAAAGT	4260
Db	4198	ACAAAGT	4257
OY	4261	GAATGTCAAGT	4320
Db	4258	GAATGTCAAGT	4317
OY	4321	CTGTCTGAATATGTGCAACACTTCAACGAGGCTTTCTGCGGTGAACAAATGAATCTCCCA	4380
Db	4318	CTGTCTGAATATGTGCAACACTTCAACGAGGCTTTCTGCGGTGAACAAATGAATCTCCCA	4377
OY	4381	GGTCTTCAGAGCAAGAGGCCAGAGAGGCTTGCACTGGAAGGGTGTGAATGAAGGTGCC	4440
Db	4378	GGTCTTCAGAGCAAGAGGCCAGAGAGGCTTGCACTGGAAGGGTGTGAATGAAGGTGCC	4437

QY	4441	AGGAATTAACAACGAGGACAGCAGAGCGTCGGACAGGAAGTACATTGTCTTGGAGGATCA	4500
Db	4438	AGGAATTAACAACGAGGACAGCAGAGCGTCGGACAGGAAGTACATTGTCTTGGAGGATCA	4497
QY	4501	AAAGTCCTCATTTATGACAAATGAAGCCGAGAACCTGGACAGAGCCGGTGGAGAAATTT	4560
Db	4498	AAAGTCCTCATTTATGACAAATGAAGCCGAGAACCTGGACAGAGCCGGTGGAGAAATTT	4557
QY	4561	GAGCTGTCCCTTCCGACGGGGATGTATCTATCATGGTCCGTGTGGTCTTCCGAATC	4620
Db	4558	GAGCTGTCCCTTCCGACGGGGATGTATCTATCATGGTCCGTGTGGTCTTCCGAATC	4617
QY	4621	GGAATTAAGCAAAAGAGATGTCCCATACTACTGAAATGGAATCTCACCCGCAATCC	4680
Db	4618	GGAATTAAGCAAAAGAGATGTCCCATACTACTGAAATGGAATCTCACCCGCAATCC	4677
QY	4681	ACCTGCTGAGCCCGGAGAACCTCTACTGTGTAGCTCCAGACTTCCCTGACAAACAGCG	4740
Db	4678	ACCTGCTGAGCCCGGAGAACCTCTACTGTGTAGCTCCAGACTTCCCTGACAAACAGCG	4737
QY	4741	TGGGTCAACCGCTTGAATCAGTTGTGACAGTGGAGAGATTCTTAGGAAAAAGCAGAA	4800
Db	4738	TGGGTCAACCGCTTGAATCAGTTGTGACAGTGGAGAGATTCTTAGGAAAAAGCAGAA	4797
QY	4801	GCTGATGCTTAACTGCTTGGAAATCCCTGCTGAAATCGAAGGTGATGACCGTCTAGAC	4860
Db	4798	GCTGATGCTTAACTGCTTGGAAATCCCTGCTGAAATCGAAGGTGATGACCGTCTAGAC	4857
QY	4861	ATGAACGTGACGCGTCCCTCAGTGCACAGGTGATGTGGTGGGCAACGAGAAAGGCTC	4920
Db	4858	ATGAACGTGACGCGTCCCTCAGTGCACAGGTGATGTGGTGGGCAACGAGAAAGGCTC	4917
QY	4921	TACGCCCTGAAATGTCTTGA AAAAATCCCTAACCATGTCCAGAAATTGGAGCAGTCTTC	4980
Db	4918	TACGCCCTGAAATGTCTTGA AAAAATCCCTAACCATGTCCAGAAATTGGAGCAGTCTTC	4977
QY	4981	CAAAATTTATATTATCAAGAGCTTGGAGAGCTACTGATATGACAGAGAAAGCGGCA	5040
Db	4978	CAAAATTTATATTATCAAGAGCTTGGAGAGCTACTGATATGACAGAGAAAGCGGCA	5037
QY	5041	CTGTGCTGTGGACGTGAAGAAAGTGAACAGTCCCTGGCCAGTCCCACTGCGCTGCC	5100
Db	5038	CTGTGCTGTGGACGTGAAGAAAGTGAACAGTCCCTGGCCAGTCCCACTGCGCTGCC	5097
QY	5101	CAGCCCGACATCTCACCCAAATTTTGAAGCTGTCAAGGGCTGCACATTGTTTGGGGCA	5160
Db	5098	CAGCCCGACATCTCACCCAAATTTTGAAGCTGTCAAGGGCTGCACATTGTTTGGGGCA	5157
QY	5161	GGCAAGATTGGAAGCGGGCTCTGTGATCTGTGCAGCCATGCCCCAGCAAGTGTCTATTCTC	5220
Db	5158	GGCAAGATTGGAAGCGGGCTCTGTGATCTGTGCAGCCATGCCCCAGCAAGTGTCTATTCTC	5217
QY	5221	CGCTACAAACGAAAACTCTACCAAAATCTGATGTCGGAAAGATAGAACCTCAGAGCCC	5280
Db	5218	CGCTACAAACGAAAACTCTACCAAAATCTGATGTCGGAAAGATAGAACCTCAGAGCCC	5277
QY	5281	TGCAGCTGTATCCACTTCAACCAATTAACAGTATCTCTATTGGAACCAATAATTTCTACGAA	5340
Db	5278	TGCAGCTGTATCCACTTCAACCAATTAACAGTATCTCTATTGGAACCAATAATTTCTACGAA	5337
QY	5341	ATGCACATGGAAGCAGTACACGCTCGAGGAATTTCTGTGATPAAGATGACCAATTCCTTGGCA	5400
Db	5338	ATGCACATGGAAGCAGTACACGCTCGAGGAATTTCTGTGATPAAGATGACCAATTCCTTGGCA	5397
QY	5401	CTGTCTGTATTTGCGCGCTCTTTCGAACAGCTTCCCTGTCTCAATCTGTGCAGGTGAACAGC	5460
Db	5398	CTGTCTGTATTTGCGCGCTCTTTCGAACAGCTTCCCTGTCTCAATCTGTGCAGGTGAACAGC	5457
QY	5461	GCAAGGGCAGGAGAGAGTACTTGTGTGTGTTTCAAGAAATTTGGAATGTTTCGAGATTC	5520
Db	5458	GCAAGGGCAGGAGAGAGTACTTGTGTGTGTTTCAAGAAATTTGGAATGTTTCGAGATTC	5517
QY	5521	TACGGAAGACGTAGCCGCAACAGCATCTCAACTGTGAGTGTGCTTACCTTTGGCTTTGCC	5580

PS Claim 20; SEQ ID NO 13; 395bp; English.

XX The invention relates to isolated NOVX polypeptides and polynucleotides.

CC NOVX polypeptides and polynucleotides are used to prevent, diagnose or

CC treat a medical condition in human related to the aberrant expression and

CC activity of NOVX polypeptides. For example, NOVX polypeptides and

CC polynucleotides may be used to treat disorders associated with decreased

CC expression or activity of NOVX by supplementing the patient with

CC production or to rectify mutations. Conversely, antisense NA molecules

CC may be administered to down regulate expression of NOVX polypeptides by

CC binding with the cells own genes and preventing their expression. NOVX

CC polynucleotides and complementary sequences may also be used as DNA

CC probes in diagnostic assays to detect and quantitate the presence of

CC similar sequences in samples, and so which patients may be in need of

CC restorative therapy. NOVX polypeptides may also be used as antigens in

CC the production of antibodies and in assays to identify modulators

CC (agonists and antagonists) of the expression and activity of NOVX. The

CC anti-NOVX polypeptide antibodies, agonists and antagonists may also be

CC used to modulate NOVX polynucleotide expression and activity of NOVX

CC polypeptides. The anti-NOVX polypeptide antibodies may also be used as

CC diagnostic agents for detecting the presence of NOVX in samples. NOVX

CC polypeptides and polynucleotides may be used in this way to prevent,

CC diagnose and treat: metabolic disorders, diabetes, obesity, infectious

CC disease, anorexia, cancer, cancer-associated cachexia, neurodegenerative

CC disorders, Alzheimer's Disease, Parkinson's Disorder, immune disorders,

CC hematopoietic disorders, and the various dyslipidaemias, metabolic

CC disturbances associated with obesity, the metabolic syndrome X and

CC wasting disorders associated with chronic diseases and various cancers.

CC They may also be used as antibacterial agents. The present sequence

CC represents DNA encoding a human NOVX protein.

XX

SO Sequence 6201 BP; 1735 A; 1551 C; 1709 G; 1206 T; 0 U; 0 Other;

Query Match 98.8%; Score 6091; DB 12; Length 6201;

Best Local Similarity 99.5%; Pred. No. 0;

Matches 6122; Conservative 0; Mismatches 30; Indels 3; Gaps 1;

QY 1 ATCTTGAAGTTCAAAATATGAGAGCGGGAATCTTTGATCTGCTGCTGAACCCATT 60

DB 1 ATCTTGAAGTTCAAAATATGAGAGCGGGAATCTTTGATCTGCTGCTGAACCCATT 60

QY 61 GCCAGCGGGCCCTCCAGGCTGAATCTGTTCCAGGGGAAACACCCCTTAAGACTCA 120

DB 61 GCCAGCGGGCCCTCCAGGCTGAATCTGTTCCAGGGGAAACACCCCTTAAGACTCA 120

QY 121 CAGCAGATGCTCTCTTCCCGAAGAGGATTAAGATCCCTCTTTGTTCTTTGAA 180

DB 121 CAGCAGATGCTCTCTTCCCGAAGAGGATTAAGATCCCTCTTTGTTCTTTGAA 180

QY 181 GAATGAGTCAGGCTGCTGTGAAGATTAAGACGTGAGCACTTTGTCGGAAGTAT 240

DB 181 GAATGAGTCAGGCTGCTGTGAAGATTAAGACGTGAGCACTTTGTCGGAAGTAT 240

QY 241 TCCGACACCAATGAGTGAAGAGGCTCCAGCTTCGCGAAGAGACTTCCAGAGTCA 300

DB 241 TCCGACACCAATGAGTGAAGAGGCTCCAGCTTCGCGAAGAGACTTCCAGAGTCA 300

QY 301 AGCTCTTGAAGTGTGTGCTCACTTGTGAAGTCAAGTGTGAAGAGAAAGCAACGGG 360

DB 301 AGCTCTTGAAGTGTGTGCTCACTTGTGAAGTCAAGTGTGAAGAGAAAGCAACGGG 360

QY 361 GACATCTATGCTATGAAGATGAAGAAAGGCTTTATTTGGCCGAGAGCAAGTTTCA 420

DB 361 GACATCTATGCTATGAAGATGAAGAAAGGCTTTATTTGGCCGAGAGCAAGTTTCA 420

QY 421 TTTTGAAGAGAGGAGCAATATATCTCGAAGCAAGCCCGTGGATCCCCCAATTA 480

DB 421 TTTTGAAGAGAGGAGCAATATATCTCGAAGCAAGCCCGTGGATCCCCCAATTA 480

QY 481 CAGTATGCTTTTCAAGACAAAATCACTTTATCTGTGATGGAATATCAAGCTGAGGG 540

DB 481 CAGTATGCTTTTCAAGACAAAATCACTTTATCTGTGATGGAATATCAAGCTGAGGG 540

QY 541 GACTTGCTGTCACTTTTGAATATGATGAGACCAATTAGATGAAAACCTGATACAGTTT 600

DB 541 GACTTGCTGTCACTTTTGAATATGATGAGACCAATTAGATGAAAACCTGATACAGTTT 600

QY 601 TACTAGCTAGCTGATTTTGGCTTCAAGGTTATCTGATGATGGATACGTGATCGA 660

DB 601 TACTAGCTAGCTGATTTTGGCTTCAAGGTTATCTGATGATGGATACGTGATCGA 660

QY 661 GAATCAAGCCTGAGAACTTCTGTTGACCGACAGACACATCAAGCTGTGATTTT 720

DB 661 GAATCAAGCCTGAGAACTTCTGTTGACCGACAGACACATCAAGCTGTGATTTT 720

QY 721 GATATGCGCGGAAATGAATTAACACAGATGATGATCCAACTCCGATTTGGAGC 780

DB 721 GATATGCGCGGAAATGAATTAACACAGATGATGATCCAACTCCGATTTGGAGC 780

QY 781 CCAATTAACATGCTCTGTAAGTCTGATGATGATGATGATGATGATGATGATGATG 840

DB 781 CCAATTAACATGCTCTGTAAGTCTGATGATGATGATGATGATGATGATGATGATG 840

QY 841 GGCCTGACTGTGACTGTGCTGATGATGATGATGATGATGATGATGATGATGATG 900

DB 841 GGCCTGACTGTGACTGTGCTGATGATGATGATGATGATGATGATGATGATGATG 900

QY 897 GGCCTGACTGTGACTGTGCTGATGATGATGATGATGATGATGATGATGATGATG 957

DB 897 GGCCTGACTGTGACTGTGCTGATGATGATGATGATGATGATGATGATGATGATG 957

QY 957 TTTTGAATTTTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1017

DB 957 TTTTGAATTTTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1017

QY 1017 TTTTGAATTTTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1077

DB 1017 TTTTGAATTTTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1077

QY 1077 TTTTGAATTTTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1137

DB 1077 TTTTGAATTTTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1137

QY 1137 TTTTGAATTTTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1197

DB 1137 TTTTGAATTTTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1197

QY 1197 TTTTGAATTTTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1257

DB 1197 TTTTGAATTTTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1257

QY 1257 TTTTGAATTTTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1317

DB 1257 TTTTGAATTTTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1317

QY 1317 TTTTGAATTTTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1377

DB 1317 TTTTGAATTTTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1377

QY 1377 TTTTGAATTTTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1437

DB 1377 TTTTGAATTTTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1437

QY 1437 TTTTGAATTTTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1497

DB 1437 TTTTGAATTTTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1497

QY 1497 TTTTGAATTTTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1557

DB 1497 TTTTGAATTTTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1557

QY 1557 TTTTGAATTTTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1617

DB 1557 TTTTGAATTTTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1617

QY 1617 TTTTGAATTTTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1677

DB 1617 TTTTGAATTTTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1677

Db 1618 GATATCAGAGAGAGAGCCGGAAGCTCCAGAAATCAAGAGAGAGGTACAGGCTCA 1677
Qy 1681 GTGGAAGAAATGAGTTGATGATGAATCAGTTGGAAGAGATCTTGTCTCAGCAAGAGA 1740
Db 1678 GTGGAAAGAAATGAGTTGATGATGAATCAGTTGGAAGAGATCTTGTCTCAGCAAGAGA 1737
Qy 1741 CGAGAGTATCTACGAATCTGAGCTGAGAGAGCTTCGGCTTGTCTGCTGAAGATTCAAG 1800
Db 1738 CGAGAGTATCTACGAATCTGAGCTGAGAGAGCTTCGGCTTGTCTGCTGAAGATTCAAG 1797
Qy 1801 CGAAAGCCGACAAATCTCAGCATTAATCTGTAAGGCTTAAGATCAAGGAAAGCCTGA 1860
Db 1798 CGAAAGCCGACAAATCTCAGCATTAATCTGTAAGGCTTAAGATCAAGGAAAGCCTGA 1857
Qy 1861 GTGGAGAAATATGCGAAATCGAGAAAGATCAATGCTGAGCAGCAGCTCAAAATTCAGAG 1920
Db 1858 GTGGAGAAATATGCGAAATCGAGAAAGATCAATGCTGAGCAGCAGCTCAAAATTCAGAG 1917
Qy 1921 CTCGAAGAGAACTGAGAGAGGCTGTAAAGCCAGCAGCAGAGGCAACCGAGCTGTGAG 1980
Db 1918 CTCGAAGAGAACTGAGAGAGGCTGTAAAGCCAGCAGCAGAGGCAACCGAGCTGTGAG 1977
Qy 1981 AATATCGCCAGGCAAGAGAGCCGAGAGGAGCTGGAAGAACTGCGAAACCGAGAG 2040
Db 1978 AATATCGCCAGGCAAGAGAGCCGAGAGGAGCTGGAAGAACTGCGAAACCGAGAG 2037
Qy 2041 GATTCTTCTGAAGGCATCAGAAAGAGCTGTGGAAGCTGAGAGCGCCGCTTCTCTG 2100
Db 2038 GATTCTTCTGAAGGCATCAGAAAGAGCTGTGGAAGCTGAGAGCGCCGCTTCTCTG 2097
Qy 2101 GAGAACAGGTAAAGAGACTAGAGACCATGAGCGTAGAGAAAAAGACTGAAGATGAC 2160
Db 2098 GAGAACAGGTAAAGAGACTAGAGACCATGAGCGTAGAGAAAAAGACTGAAGATGAC 2157
Qy 2161 ATCCAGCAAAATCCCAACGATTCAGCAGATGCTGATTAATTTCTGAGCTCGAAGAG 2220
Db 2158 ATCCAGCAAAATCCCAACGATTCAGCAGATGCTGATTAATTTCTGAGCTCGAAGAG 2217
Qy 2221 AAACATCGGAGGCGCCAAATCTCAGCCAGCAGACCTAGAGTGCACCTGAAACGAAAGAG 2280
Db 2218 AAACATCGGAGGCGCCAAATCTCAGCCAGCAGACCTAGAGTGCACCTGAAACGAAAGAG 2277
Qy 2281 CAGCACTATGAGAAAAAGATTAAAGTTTGAACATACATGATTAAGAAAGACTGTGAC 2340
Db 2278 CAGCACTATGAGAAAAAGATTAAAGTTTGAACATACATGATTAAGAAAGACTGTGAC 2337
Qy 2341 AAGGAGCACTGAGAAACATGATGAGAGACAGAGAGAGAGGCCCTGAGAAAGGCAAA 2400
Db 2338 AAGGAGCACTGAGAAACATGATGAGAGACAGAGAGAGAGGCCCTGAGAAAGGCAAA 2397
Qy 2401 ATTCTCAGCGAACAAGAGGAGGATGATCAATGCTATGATTCGAAGTCCAAATCCCTGAA 2460
Db 2398 ATTCTCAGCGAACAAGAGGAGGATGATCAATGCTATGATTCGAAGTCCAAATCCCTGAA 2457
Qy 2461 CAGAGATTTGTGAACTGTCTGAAGCCAAATACTTCAGCAAAATAGCAGCTTTTACC 2520
Db 2458 CAGAGATTTGTGAACTGTCTGAAGCCAAATACTTCAGCAAAATAGCAGCTTTTACC 2517
Qy 2521 CAAAGGAACTGAGAGGCCCAAGAGAGATGATTTCTGAACTCAGGCAAGAAATTTTAC 2580
Db 2518 CAAAGGAACTGAGAGGCCCAAGAGAGATGATTTCTGAACTCAGGCAAGAAATTTTAC 2577
Qy 2581 CTGGAACACAGGCTGGAGAGTTGAGAGCCCAAGAACGAAACTGAGAGAGCAGCTGGAG 2640
Db 2578 CTGGAACACAGGCTGGAGAGTTGAGAGCCCAAGAACGAAACTGAGAGAGCAGCTGGAG 2637
Qy 2641 AAGATCAGCCACCAAGACCAAGTGAACAAGATCGCTGCTGGAACCTGAGACCAAGATTG 2700
Db 2638 AAGATCAGCCACCAAGACCAAGTGAACAAGATCGCTGCTGGAACCTGAGACCAAGATTG 2697
Qy 2701 CGGAGGTCAGTCTTAGACAAGAGAGCAAGAACTGAGGCTCAAGCCGCTCAAGAG 2760

Db 2698 CGGAGGTGAGTCTTAGACACAGAGAGCAAGAACTGAGGCTCAAGCCGCTCAAGAG 2757
Qy 2761 CTACAGCTCTCCCTGAGAGAGCCGAGTCAAGATTGAACAGCCCTGACAGCTGACGGCG 2820
Db 2758 CTACAGCTCTCCCTGAGAGAGCCGAGTCAAGATTGAACAGCCCTGACAGCTGACGGCG 2817
Qy 2821 GCCCTGAGAGCCAGCTTGCCAGGCGAAGACAGAGCTGGAAGAGACCAAGCAAGAGCT 2880
Db 2818 GCCCTGAGAGCCAGCTTGCCAGGCGAAGACAGAGCTGGAAGAGACCAAGCAAGAGCT 2877
Qy 2881 GAAGAGAGATCCAGGCACTCAGGCAATAGAGATGAATCCAGGCAAAATTTGATGCT 2940
Db 2878 GAAGAGAGATCCAGGCACTCAGGCAATAGAGATGAATCCAGGCAAAATTTGATGCT 2937
Qy 2941 CTTCTGTAACAGCTGTACTGTATACAGACCTGAGAGAGCAGCTTAAACAGCTGACCGAG 3000
Db 2938 CTTCTGTAACAGCTGTACTGTATACAGACCTGAGAGAGCAGCTTAAACAGCTGACCGAG 2997
Qy 3001 GACACGCTGAATCTCAACAAACCAAACTTCTACTTGTCCAAACAATCGATGAGGCTTCT 3060
Db 2998 GACACGCTGAATCTCAACAAACCAAACTTCTACTTGTCCAAACAATCGATGAGGCTTCT 3057
Qy 3061 GGGCGCAACGAGAGATTGTACAACTGCCAGAGTGAAGTGAACATCTCCGCGGGAGATC 3120
Db 3058 GGGCGCAACGAGAGATTGTACAACTGCCAGAGTGAAGTGAACATCTCCGCGGGAGATC 3117
Qy 3121 ACGGAACGAGAGTGAAGCTTACACAGCAGAAACAAACGATGAGGCTGTGAAGACCAAG 3180
Db 3118 ACGGAACGAGAGTGAAGCTTACACAGCAGAAACAAACGATGAGGCTGTGAAGACCAAG 3177
Qy 3178 TGCAACATGCTGAGAGAAACAGTCAATGATTTGAGAGCCCTTAAACGATGAGCTTGA 3237
Db 3181 TGCAACATGCTGAGAGAAACAGTCAATGATTTGAGAGCCCTTAAACGATGAGCTTGA 3240
Qy 3241 AAAGACCGCAGTGGAGAGCCTGGAAGAGAGCTGAGTGAAGTGAAGTGAAGTGAAGTGAAG 3300
Db 3238 AAAGACCGCAGTGGAGAGCCTGGAAGAGAGCTGAGTGAAGTGAAGTGAAGTGAAGTGAAG 3297
Qy 3301 TGTCCGATTCGAGAGCTGCAAGAGATGCTGGAACCCGAGAAACAGAGAGGCGAGAGCC 3360
Db 3298 TGTCCGATTCGAGAGCTGCAAGAGATGCTGGAACCCGAGAAACAGAGAGGCGAGAGCC 3357
Qy 3361 GATCAGCGATCAACGAGTCTGCGCAGGTGTGAGAGCTGCAAGTGAAGAGACCAAGGCT 3420
Db 3358 GATCAGCGATCAACGAGTCTGCGCAGGTGTGAGAGCTGCAAGTGAAGAGACCAAGGCT 3417
Qy 3421 GAGATTTCTGCTTGCAGCAGAGCTCTCAAGAGACAAAGCTGAAGGCCGAGAGCTTCT 3480
Db 3418 GAGATTTCTGCTTGCAGCAGAGCTCTCAAGAGACAAAGCTGAAGGCCGAGAGCTTCT 3477
Qy 3481 GACAGCTCAATGACCTGGAGAAAGACATGTATGCTTGAATGATGATGCCGAGAGCTTA 3540
Db 3478 GACAGCTCAATGACCTGGAGAAAGACATGTATGCTTGAATGATGATGCCGAGAGCTTA 3537
Qy 3541 CAGCAGAGCTGAGACTGAACGAGAGCTCAACAGAGGCTTCTGGAAGAGCAAGCCAA 3600
Db 3538 CAGCAGAGCTGAGACTGAACGAGAGCTCAACAGAGGCTTCTGGAAGAGCAAGCCAA 3597
Qy 3601 TTACAGACAGATGAGACTGCAAGAAATCACATTTTCCGTCTGACTCAAGACTGCA 3660
Db 3598 TTACAGACAGATGAGACTGCAAGAAATCACATTTTCCGTCTGACTCAAGACTGCA 3657
Qy 3661 GAAGCTCTAGATGGGCTGATCTACTGAAGACGAAAGAGTGAAGCTTGGAGTATCAGCTG 3720
Db 3658 GAAGCTCTAGATGGGCTGATCTACTGAAGACGAAAGAGTGAAGCTTGGAGTATCAGCTG 3717
Qy 3721 GAAACATTCAAGTTCTTATTTCTCATGAAAGGTGAAGTGAAGCACTATTTCTCAA 3780
Db 3718 GAAACATTCAAGTTCTTATTTCTCATGAAAGGTGAAGTGAAGCACTATTTCTCAA 3777
Qy 3781 CAAACCAAACTATTGATTTTCTGCAAGCCAAATGGAACCAACTGTCTAAAAAGAAAAAG 3840
Db 3778 CAAACCAAACTATTGATTTTCTGCAAGCCAAATGGAACCAACTGTCTAAAAAGAAAAAG 3837

QY 3841 GTTCTCTGAGTAACAATGAGCTGAAGCTGGCCCTGAGAAAGAAAGCTCGCTGCA 3900
| | | | |
Db 3838 GTGCTCTGAGTAACAATGAGCTGAAGCTGGCCCTGAGAAAGAAAGCTCGCTGCA 3897
QY 3901 GAGCTAGAGAAAGCCCTTACAGAAAGCCGATCGAGCTCCGGTCCGGCCGGAGAAAGCT 3960
| | | | |
Db 3898 GAGCTAGAGAAAGCCCTTACAGAAAGCCGATCGAGCTCCGGTCCGGCCGGAGAAAGCT 3957
QY 3961 GCCCACCAGAAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 4020
| | | | |
Db 3958 GCCCACCAGAAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 4017
QY 4021 ATGGCAATGCTCCGCAATCGAGGCTCCGCAAGAGCAAGCAAGCAAGCAAGCAAGCAAG 4080
| | | | |
Db 4018 ATGGCAATGCTCCGCAATCGAGGCTCCGCAAGAGCAAGCAAGCAAGCAAGCAAGCAAG 4077
QY 4081 GCCCCGCAATCGAGGCTCCGCAAGAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 4140
| | | | |
Db 4078 GCCCCGCAATCGAGGCTCCGCAAGAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 4137
QY 4141 AAGGAACGATGACCAACAATATCTTCAACGATTCACGATTCACGATTCACGATTCACG 4200
| | | | |
Db 4138 AAGGAACGATGACCAACAATATCTTCAACGATTCACGATTCACGATTCACGATTCACG 4197
QY 4201 ACAAGGTGCTGTGTCTGTGATACCTGTGACCTTTGACGCGACGACGACGACGACGAC 4260
| | | | |
Db 4198 ACAAGGTGCTGTGTCTGTGATACCTGTGACCTTTGACGCGACGACGACGACGACGAC 4257
QY 4261 GAATGTCAAGGTGATGTGTCAACCCCAAGTGTCCACGCTTGTGACGACGACGACGACG 4320
| | | | |
Db 4258 GAATGTCAAGGTGATGTGTCAACCCCAAGTGTCCACGCTTGTGACGACGACGACGACG 4317
QY 4321 CCTGCTGAATATGCAACAATTACCGAGGCTTCTGCGGTGACAAATGAATCTCCCA 4380
| | | | |
Db 4318 CCTGCTGAATATGCAACAATTACCGAGGCTTCTGCGGTGACAAATGAATCTCCCA 4377
QY 4381 GGTTCTCAAGCCCAAGAGCCCAAGACGATTTGACCTGGAAGGTGATGAAGGTGCCC 4440
| | | | |
Db 4378 GGTTCTCAAGCCCAAGAGCCCAAGACGATTTGACCTGGAAGGTGATGAAGGTGCCC 4437
QY 4441 AAGGAATACAAAGCAAGCAAGCAAGGCTGGGACAGGAAGTCAATTTGCTCGAGGGATCA 4500
| | | | |
Db 4438 AAGGAATACAAAGCAAGCAAGCAAGGCTGGGACAGGAAGTCAATTTGCTCGAGGGATCA 4497
QY 4501 AAAGTCTCATTTATGACATGAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 4560
| | | | |
Db 4498 AAAGTCTCATTTATGACATGAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 4557
QY 4561 GAGCTGTGCTTCCGACCGGAGATGTATCTATTCATGTGTGCGTTGGTCTTCCGAATC 4620
| | | | |
Db 4558 GAGCTGTGCTTCCGACCGGAGATGTATCTATTCATGTGTGCGTTGGTCTTCCGAATC 4617
QY 4621 GCAAAATACAGCCAAAGCAAGTGTCCCATACATCTGAAGTGAATCTGACCCGACACC 4680
| | | | |
Db 4618 GCAAAATACAGCCAAAGCAAGTGTCCCATACATCTGAAGTGAATCTGACCCGACACC 4677
QY 4681 ACCTGTGACCGGAGAAACCTCTATCTTGTCTAGCTCCAGCTTCCCTGACAAACAGCC 4740
| | | | |
Db 4678 ACCTGTGACCGGAGAAACCTCTATCTTGTCTAGCTCCAGCTTCCCTGACAAACAGCC 4737
QY 4741 TGGGTACCGGCTTGAATCAAGTTGTGCGAGGTGGAGAGTTCTAGGGAAAAAGCAGAA 4800
| | | | |
Db 4738 TGGGTACCGGCTTGAATCAAGTTGTGCGAGGTGGAGAGTTCTAGGGAAAAAGCAGAA 4797
QY 4801 GCTGATGCTAAACTGCTTGAAGAACTCCGCTGAAGCTGAAGGTGAATGACGCTCAGAC 4860
| | | | |
Db 4798 GCTGATGCTAAACTGCTTGAAGAACTCCGCTGAAGCTGAAGGTGAATGACGCTCAGAC 4857
QY 4861 ATGAACCTGACGCTGCTTCAAGTGAACAGTGTGTGTGTGGGACCCAGAGAGGGCTC 4920
| | | | |
Db 4858 ATGAACCTGACGCTGCTTCAAGTGAACAGTGTGTGTGTGGGACCCAGAGAGGGGCT 4917

QY 4921 TACGCCCTGAATGTCTTGAAGAACTCCCTTAACCATGTCCAGAGAAATTGAGCAGTCTTC 4980
| | | | |
Db 4918 CTACGCTGATGTCTTGAAGAACTCCCTTAACCATGTCCAGAGAAATTGAGCAGTCTTC 4977
QY 4981 CAATTTTATTTATTCAGAGACCTGAGAGCTACTCATGTATGACAGAGAAAGCGGGCA 5040
| | | | |
Db 4978 CAATTTTATTTATTCAGAGACCTGAGAGACCTACTCATGTATGACAGAGAAAGCGGGCA 5037
QY 5041 CTGTGTCTGTGACGCTGAAGAAAGTGAACAGTCCCTGGCCAGTCCCACTGCTGCC 5100
| | | | |
Db 5038 CTGTGTCTGTGACGCTGAAGAAAGTGAACAGTCCCTGGCCAGTCCCACTGCTGCC 5097
QY 5101 CAGCCGCAATCTCACCCCAACATTTTGAAGCTGTCAAGGGCTGCCACTTGTGTGGGCA 5160
| | | | |
Db 5098 CAGCCGCAATCTCACCCCAACATTTTGAAGCTGTCAAGGGCTGCCACTTGTGTGGGCA 5157
QY 5161 GGCAGATTTGAGAACGGGCTCTGATCTGTGACGACATGCGCAAGGCTGATCTC 5220
| | | | |
Db 5158 GGCAGATTTGAGAACGGGCTCTGATCTGTGACGACATGCGCAAGGCTGATCTC 5217
QY 5221 CGCTACAGAGAAACCTCAGCAATATCTGATCCGAGAAAGATAGACCTCAGAGCC 5280
| | | | |
Db 5218 CGCTACAGAGAAACCTCAGCAATATCTGATCCGAGAAAGATAGACCTCAGAGCC 5277
QY 5281 TGCAGCTGTATCTCATCTTCAACAAATTACATATCTTATGAAACCAATAATTCTACGA 5340
| | | | |
Db 5278 TGCAGCTGTATCTCATCTTCAACAAATTACATATCTTATGAAACCAATAATTCTACGA 5337
QY 5341 ATGACATGAAGAGATGACAGCTCGAGGAATTCGATTAAGATGAACCAATTCCTTGGCA 5400
| | | | |
Db 5338 ATGACATGAAGAGATGACAGCTCGAGGAATTCGATTAAGATGAACCAATTCCTTGGCA 5397
QY 5401 CCTGCTGTGTGTGCGGCTCTTTCACACAGCTTCCCTGTCTCAATCGTGAAGTGAACAG 5460
| | | | |
Db 5398 CCTGCTGTGTGTGCGGCTCTTTCACACAGCTTCCCTGTCTCAATCGTGAAGTGAACAG 5457
QY 5461 GCAGGGCAGCAGAGAGATCTTGTGTGTTCACGAATTTGGAATGTTGTGGAATTTCT 5520
| | | | |
Db 5458 GCAGGGCAGCAGAGAGATCTTGTGTGTTCACGAATTTGGAATGTTGTGGAATTTCT 5517
QY 5521 TACGGAAGCTGAGCGCAGACAGATCTCAATGAGTGCCTTATCTTGGCTTGGC 5580
| | | | |
Db 5518 TACGGAAGCTGAGCGCAGACAGATCTCAATGAGTGCCTTATCTTGGCTTGGC 5577
QY 5581 TACAGAGAACCTTATCTTGTGTGACCACTTCAACTCACTGAAATTAATTGATCCAG 5640
| | | | |
Db 5578 TACAGAGAACCTTATCTTGTGTGACCACTTCAACTCACTGAAATTAATTGATCCAG 5637
QY 5641 GCAAGCTCTCAGCAGAGAACCTTGTGCGAGGTACCTGGAATCCGAAACCGCGCTAC 5700
| | | | |
Db 5638 GCAAGCTCTCAGCAGAGAACCTTGTGCGAGGTACCTGGAATCCGAAACCGCGCTAC 5697
QY 5701 CTGGGCGCTGCAATTTCTCAGAGAGATTTTACTTGGAGTCTCATACAGGATTAATTA 5760
| | | | |
Db 5698 CTGGGCGCTGCAATTTCTCAGAGAGATTTTACTTGGAGTCTCATACAGGATTAATTA 5757
QY 5761 AAGGTCAATTTGTGACAGAGAAACCTGTGAAGAGTCCGAGCTGAACCAACCGGGC 5820
| | | | |
Db 5758 AAGGTCAATTTGTGACAGAGAAACCTGTGAAGAGTCCGAGCTGAACCAACCGGGC 5817
QY 5821 CCGTCACTCCCGACAGACGCCCAACAGAGAGGCCCACTCATGACAGACATC 5880
| | | | |
Db 5818 CCGTCACTCCCGACAGACGCCCAACAGAGAGGCCCACTCATGACAGACATC 5877
QY 5881 ACCAAGCGGTGGCTTCAAGCCAGCGCGCGCCGAAGGCCCAAGCCAGCGAGAGCA 5940
| | | | |
Db 5878 ACCAAGCGGTGGCTTCAAGCCAGCGCGCGCCGAAGGCCCAAGCCAGCGAGAGCA 5937
QY 5941 AGCACAACCCACCGCTTACCGAGAGGGGCGGACCGAGCTGCGAGAGGACAAAGTCTCTGGC 6000
| | | | |
Db 5938 AGCACAACCCACCGCTTACCGAGAGGGGCGGACCGAGCTGCGAGAGGACAAAGTCTCTGGC 5997
QY 6001 CGCCCTTGAAGCGAGAAAGTCCCCCGAGCGGATTACTACGACGCGAGAGAGCGGTCC 6060

Db 5998 CGCCCCCTGAGCGAGAAATCCCGCGCGATGCTCAGACGCGAGAGAGCGGATCC 6057
Qy 6061 CCCGCGAGGCTGTTTGAAGACAGACAGAGGGGCGGCTGCTCGGGAGACCGTAAAGACC 6120
Db 6058 CCGCGAGGCTGTTTGAAGACAGACAGAGGGGCGGCTGCTCGGGAGACCGTAAAGACC 6117
Qy 6121 CCGCTGTCCAGGTGAACAGGTCTGGAGCCAGTC 6155
Db 6118 CCGCTGTCCAGGTGAACAGGTGAGGACGATTC 6152
RESULT 14
AB07871
ID AB07871 standard; cDNA; 5877 BP.
XX AB07871;
AC AB07871;
XX 10-OCT-2002 (first entry)
XX Human kinase cDNA #2.
XX
XX Human; kinase; enzyme; serine-threonine kinase; motropic; cytosolic;
KM Citron rho-interacting kinase; gene therapy; mental disorder; cancer;
KM gene; ss.
XX Homo sapiens.
XX
XX Key Location/Qualifiers
FT 1..5877
FT CDS /*tag= a
PT /product= "Kinase"
PN W0200259325-A2.
XX 01-AUG-2002.
XX 20-DEC-2001; 2001WO-US050497.
XX 27-DEC-2000; 2000US-0258335P.
XX (LEXI-) LEXICON GENETICS INC.
XX Yu X, Miranda M, Fiddle CJ;
PI WPI: 2002-599796/64.
XX P-PSDB; ABB81928.
XX Novel polynucleotide encoding human proteins that are structurally
PT similar to animal kinases, useful for drug screening, diagnosis, in gene
PT therapy of disorders and diseases e.g. cancer and pharmacogenomic
PT applications.
XX Disclosure; Page 44-45; 50pp; English.
XX
XX The invention relates to a novel human protein that shares structural
CC similarity with animal kinases, including serine-threonine kinases,
CC particularly Citron rho-interacting kinases. The proteins of the
CC invention have motropic and cytosolic activity. The polynucleotides may
CC have a use in gene therapy. The encoded novel polypeptides are useful for
CC generating antibodies, as reagents in diagnostic assays, for identifying
CC other cellular gene products related to NHP and as reagents in assays for
CC screening for compounds that are useful in the treatment of mental,
CC biological or medical disorders and diseases including cancer. The
CC sequence encodes a novel human kinase of the invention
XX
XX Sequence 5877 BP; 1680 A; 1433 C; 1583 G; 1181 T; 0 U; 0 Other;
SQ
Query Match 94.7%; Score 5837; DB 6; Length 5877;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 5837; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ATGTGAAGTTCAATATGAGCGCGAATCTTTGATGCTGTGCTGAACCCATT 60

Db 1 ATGTGAAGTTCAATATGAGCGCGAATCTTTGATGCTGTGCTGAACCCATT 60
Qy 61 GCCAGCGGCGCTCCAGGCTGAATCTGTTCTCCAGGGGAAACCAACCTTTATAGACTCA 120
Db 61 GCCAGCGGCGCTCCAGGCTGAATCTGTTCTCCAGGGGAAACCAACCTTTATAGACTCA 120
Qy 121 CAGCAGATGCTCTCTTTCCCGAAGAGATATAGATGCTCTTTGTTCTTTGAA 180
Db 121 CAGCAGATGCTCTCTTTCCCGAAGAGATATAGATGCTCTTTGTTCTTTGAA 180
Qy 181 GAATGAGTACGCTGCTGATGATGATTAAGATGAGCACTTTGCTCGGAAGAT 240
Db 181 GAATGAGTACGCTGCTGATGATGATTAAGATGAGCACTTTGCTCGGAAGAT 240
Qy 241 TCCGACACATAGCTGATGATTAAGATGAGCACTTTGCTCGGAAGAT 300
Db 241 TCCGACACATAGCTGATGATTAAGATGAGCACTTTGCTCGGAAGAT 300
Qy 301 AGTCTGTAGGTTGTGTGCTGATGATGATGAGTGAAGTGAAGAAAGCAACCGG 360
Db 301 AGTCTGTAGGTTGTGTGCTGATGATGATGAGTGAAGTGAAGAAAGCAACCGG 360
Qy 361 GACATCTATGCTATGAAAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 420
Db 361 GACATCTATGCTATGAAAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 420
Qy 421 TTTTGTGAGAGAGCGGAAATATATATCTGAGAGCAAGCCGCTGATCCCAATTA 480
Db 421 TTTTGTGAGAGAGCGGAAATATATATCTGAGAGCAAGCCGCTGATCCCAATTA 480
Qy 481 CAGTATGCTTTTGAAGCAAAATATCACTTTATCTGTGCTATGAAATATCAAGCTGAG 540
Db 481 CAGTATGCTTTTGAAGCAAAATATCACTTTATCTGTGCTATGAAATATCAAGCTGAG 540
Qy 541 GACTTGCTGCTGCTTTTGAATATATATGAGAGCAAGTATGAAATATCAAGCTGAG 600
Db 541 GACTTGCTGCTGCTTTTGAATATATATGAGAGCAAGTATGAAATATCAAGCTGAG 600
Qy 601 TACCTAGCTGATGATTTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660
Db 601 TACCTAGCTGATGATTTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660
Qy 661 GACATCAAGCTGAGCAATTTCTGTTGACCGAGACAGACATCAAGCTGAGATTTT 720
Db 661 GACATCAAGCTGAGCAATTTCTGTTGACCGAGACAGACATCAAGCTGAGATTTT 720
Qy 721 GATCTGCGCGAAATGAAATTCACACAGATGATGATGATGATGATGATGATGATG 780
Db 721 GATCTGCGCGAAATGAAATTCACACAGATGATGATGATGATGATGATGATGATG 780
Qy 781 CCAATTCAGATGCTGCTGAGAGCTGACCTGATGAGAGAGAGAGAGAGAGAGAGAG 840
Db 781 CCAATTCAGATGCTGCTGAGAGCTGACCTGATGAGAGAGAGAGAGAGAGAGAGAG 840
Qy 841 GGCCTGACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 900
Db 841 GGCCTGACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 900
Qy 901 TCCCTGCGAG 960
Db 901 TCCCTGCGAG 960
Qy 961 TTTTGAATTTTCAAGTGAACCCCAAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAG 1020
Db 961 TTTTGAATTTTCAAGTGAACCCCAAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAG 1020
Qy 1021 TTGTGTGCGCGAG 1080
Db 1021 TTGTGTGCGCGAG 1080
Qy 1081 TCTAAATTAAGTGAAG 1140

Dh 1081 TCTAAAATTGACTGGAAACAATTCGTAACTCTCTCCCCCTTCGTTCCACCCCTCAG 1140
Qy 1141 TCTGACGATGACCTTCCAAATTTTGTATGTAACCAAGAAATTGCTGGTTTCATCTCT 1200
Db 1141 TCTACCAATGACACTCCAAATTTTGTATGTAACCAAGAAATTGCTGGTTTCATCTCT 1200
Qy 1201 CCGTGCCAGCTGAGCCCCCTCAGGCTTCCTCGGTTGAAGAACTGCCGTTTGTGGGGTTTTCG 1260
Db 1201 CCGTGCCAGCTGAGCCCCCTCAGGCTTCCTCGGTTGAAGAACTGCCGTTTGTGGGGTTTTCG 1260
Qy 1261 TACAGCAAGGCACTGCGGGATTCTTGTAGATCTGAGTCTGTTGTGTGCTGGGCTCTGACTCC 1320
Db 1261 TACAGCAAGGCACTGCGGGATTCTTGTAGATCTGAGTCTGTTGTGTGCTGGGCTCTGACTCC 1320
Qy 1321 CTTGCCAAGACTACTCCATCTGAAAAGAACTTTCATCAAAAGCAAAAGACTCAAGAC 1380
Db 1321 CTTGCCAAGACTACTCCATCTGAAAAGAACTTTCATCAAAAGCAAAAGACTCAAGAC 1380
Qy 1381 TCTCAGGACAAGTGTCAACAAGTAGAGCAAGAAATGACCCTGTTACATCCGAGTGTCA 1440
Db 1381 TCTCAGGACAAGTGTCAACAAGTAGAGCAAGAAATGACCCTGTTACATCCGAGTGTCA 1440
Qy 1441 GAGGTGAGGCTGTGTTAGTCAAGAAAGAGTGGAGCTGAAGGCTCTGAGACTCAGAGA 1500
Db 1441 GAGGTGAGGCTGTGTTAGTCAAGAAAGAGTGGAGCTGAAGGCTCTGAGACTCAGAGA 1500
Qy 1501 TCCCTCTGAGACGAGACTTCTGTACTTCAATCAACAGATGCACTAGACTTAAAGCAAGT 1560
Db 1501 TCCCTCTGAGACGAGACTTCTGTACTTCAATCAACAGATGCACTAGACTTAAAGCAAGT 1560
Qy 1561 TTGGAGCAAGCAAGGATGAGGTGTCCAGAGAGTAGACAAGCACTGCAAGCTTTCAT 1620
Db 1561 TTGGAGCAAGCAAGGATGAGGTGTCCAGAGAGTAGACAAGCACTGCAAGCTTTCAT 1620
Qy 1621 GATATCAGAGACGAGCCGGAAGCTTCAAGAAATCAAGAGCAGAGTCAAGGCTCAA 1680
Db 1621 GATATCAGAGACGAGCCGGAAGCTTCAAGAAATCAAGAGCAGAGTCAAGGCTCAA 1680
Qy 1681 GTGGAAGAAATGAGTGTGATGATGATCACTTGGAGAGAGATCTTGTCTCAGCAAGAGA 1740
Db 1681 GTGGAAGAAATGAGTGTGATGATGATCACTTGGAGAGAGATCTTGTCTCAGCAAGAGA 1740
Qy 1741 CGGAGTATCTTACGAATGTGAGTGTGAGAGTCTCGGCTTCTGCTGAAGAAATTCAG 1800
Db 1741 CGGAGTATCTTACGAATGTGAGTGTGAGAGTCTCGGCTTCTGCTGAAGAAATTCAG 1800
Qy 1801 CGGAAAGCGACAGAAATGTAGCATTAATCTTGAAGGCTAAGATCAAGGAAAGCTGAA 1860
Db 1801 CGGAAAGCGACAGAAATGTAGCATTAATCTTGAAGGCTAAGATCAAGGAAAGCTGAA 1860
Qy 1861 GTGGGAGATATCGAAATGTGAGAAAGATCAATGTGAGCAGACTCAAAATTCAGAG 1920
Db 1861 GTGGGAGATATCGAAATGTGAGAAAGATCAATGTGAGCAGACTCAAAATTCAGAG 1920
Qy 1921 CTCAGAAGAACTGAGAAAGGCTGTAAAAAGCCAGACGAGGCTCAAGGCTGCTCAG 1980
Db 1921 CTCAGAAGAACTGAGAAAGGCTGTAAAAAGCCAGACGAGGCTCAAGGCTGCTCAG 1980
Qy 1981 AATATCCGCGCAGCAAGAGCGAGCCGAGAGGAGCTGGAGAAAGCTGCGAAGCCAGAG 2040
Db 1981 AATATCCGCGCAGCAAGAGCGAGCCGAGAGGAGCTGGAGAAAGCTGCGAAGCCAGAG 2040
Qy 2041 GATTCTTCTGAAAGCATCAAGAAAGAACTGTGTGAAGCTGAGGAAAGCCGCTTCTCG 2100
Db 2041 GATTCTTCTGAAAGCATCAAGAAAGAACTGTGTGAAGCTGAGGAAAGCCGCTTCTCG 2100
Qy 2101 GAGAACAAAGTAAAGAGACTGAGAACATGAGAGCTGAGAGAAACAGACTGAAAGATGAC 2160
Db 2101 GAGAACAAAGTAAAGAGACTGAGAACATGAGAGCTGAGAGAAACAGACTGAAAGATGAC 2160
Qy 2161 ATCCAGACAAATCCCAAGAGATCCAGAGATGCTGATAAATTTCTGAGCTCGAAGAG 2220
Db 2161 ATCCAGACAAATCCCAAGAGATCCAGAGATGCTGATAAATTTCTGAGCTCGAAGAG 2220

Qy 2221 AAACATCGGAGAGCCCAAGTCTCAGCCAGCACTTAAGTGCACCTGAAAAGAG 2280
Db 2221 AAACATCGGAGAGCCCAAGTCTCAGCCAGCACTTAAGTGCACCTGAAAAGAG 2280
Qy 2281 CAGCACTATGAGGAAAGATTTAAAGTGTGAGCAATGATTAAGAAAGACCTGGCTGAC 2340
Db 2281 CAGCACTATGAGGAAAGATTTAAAGTGTGAGCAATGATTAAGAAAGACCTGGCTGAC 2340
Qy 2341 AAGGAGACACTGAGAAACATGATCAGAGACACAGAGAGAGGCCCATGAGAGGCAAA 2400
Db 2341 AAGGAGACACTGAGAAACATGATCAGAGACACAGAGAGAGGCCCATGAGAGGCAAA 2400
Qy 2401 ATTCTCAGCAACAGAAAGCGATGATCAATGCTATGAGATTCCAAGATCAATCCCTGAA 2460
Db 2401 ATTCTCAGCAACAGAAAGCGATGATCAATGCTATGAGATTCCAAGATCAATCCCTGAA 2460
Qy 2461 CAGAGATTGTGGAATCTGTGAAAGCCAAATTAACCTGCACAAATAGACTTTTAAAC 2520
Db 2461 CAGAGATTGTGGAATCTGTGAAAGCCAAATTAACCTGCACAAATAGACTTTTAAAC 2520
Qy 2521 CAAGGAACTATGAAGGCCCAAGAGAGATGATTTCTGAACTCAAGGCAAGAAATTTAC 2580
Db 2521 CAAGGAACTATGAAGGCCCAAGAGAGATGATTTCTGAACTCAAGGCAAGAAATTTAC 2580
Qy 2581 CTGAGACACAGGCTGGGAAATTGAGAGCCAGAAACCGAAACTGAGAGAGCTTGAG 2640
Db 2581 CTGAGACACAGGCTGGGAAATTGAGAGCCAGAAACCGAAACTGAGAGAGCTTGAG 2640
Qy 2641 AAGATCAACCAACCAACCAACCACTGACCAAGAAATCGGCTGTGGAATGAGAGCAAGTTG 2700
Db 2641 AAGATCAACCAACCAACCAACCACTGACCAAGAAATCGGCTGTGGAATGAGAGCAAGTTG 2700
Qy 2701 CGGAGAGTCAGTCTAAGACACAGAGACAGAACTGAGAGCTCAAGGCTCAAGAG 2760
Db 2701 CGGAGAGTCAGTCTAAGACACAGAGACAGAACTGAGAGCTCAAGGCTCAAGAG 2760
Qy 2761 CTACAGCTTCCCTCAGAGAGCCGAGTCAAGTTGACAGCCCTGCAAGGCTGACGAGCG 2820
Db 2761 CTACAGCTTCCCTCAGAGAGCCGAGTCAAGTTGACAGCCCTGCAAGGCTGACGAGCG 2820
Qy 2821 GCCCTGAGAGCAGCTTGGCCAGAGGAGACAGACTGGAAGACCAAGAGAAAGT 2880
Db 2821 GCCCTGAGAGCAGCTTGGCCAGAGGAGACAGACTGGAAGACCAAGAGAAAGT 2880
Qy 2881 GAAAGAGATCCAGGCACTCAGGCACTACGACATAGAGTGAATCCAGGCAAAATTTGATGCT 2940
Db 2881 GAAAGAGATCCAGGCACTCAGGCACTACGAGCAATAGAGTGAATCCAGGCAAAATTTGATGCT 2940
Qy 2941 CTTTGTACAGCTGTACTGTAACTACAGACTGTGAGAGCAGCTAAACAGCTGACCGAG 3000
Db 2941 CTTTGTACAGCTGTACTGTAACTACAGACTGTGAGAGCAGCTAAACAGCTGACCGAG 3000
Qy 3001 GACAAAGCTGAATCTCAACCAACCAAACTTCTACTTGTCCAAACACTCGATGAGGCTTCT 3060
Db 3001 GACAAAGCTGAATCTCAACCAACCAAACTTCTACTTGTCCAAACACTCGATGAGGCTTCT 3060
Qy 3061 GCGCGCAACGACGAGATTGTACAACTGCGAAGTGAAGTGAACATCTCCGCGGAGATC 3120
Db 3061 GCGCGCAACGACGAGATTGTACAACTGCGAAGTGAAGTGAACATCTCCGCGGAGATC 3120
Qy 3121 ACGGAAACGAGATGAGCTTACAGCCAGAAAGCAACGATGAGGCTTGAAGACCAAG 3180
Db 3121 ACGGAAACGAGATGAGCTTACAGCCAGAAAGCAACGATGAGGCTTGAAGACCAAG 3180
Qy 3181 TGCACCATGTCTGAGAGAAAGATCATGATTTGAGAGCCCTTAAAGATGAGTCTGTAGAA 3240
Db 3181 TGCACCATGTCTGAGAGAAAGATCATGATTTGAGAGCCCTTAAAGATGAGTCTGTAGAA 3240
Qy 3241 AAAGAGCGGAGTGGAGGCTCGAGAGAGCGTCTGGGTTGAGAGAAATCCAGTTTGAAG 3300
Db 3241 AAAGAGCGGAGTGGAGGCTCGAGAGAGCGTCTGGGTTGAGAGAAATCCAGTTTGAAG 3300

QY 3301 TGTGCGTTTGAAGAGCTGCAGAGATGCTGACACCGAGAAACAGAGACGAGGCGAGAGCC 3360
DB 3301 TGTGCGTTTGAAGAGCTGCAGAGATGCTGACACCGAGAAACAGAGACGAGGCGAGAGCC 3360
QY 3361 GATCAGCGGATCAACGAGTCTGCGCAGAGTGGAGTGGCAGAGGAGAGACACAGAGCT 3420
DB 3361 GATCAGCGGATCAACGAGTCTGCGCAGAGTGGAGTGGCAGAGGAGAGACACAGAGCT 3420
QY 3421 GAGATTCTGCTCTGCAGCAGAGCTCTCAAGAGCAGAGAGCTGAAGGCGGAGAGCTCTCT 3480
DB 3421 GAGATTCTGCTCTGCAGCAGAGCTCTCAAGAGCAGAGAGCTGAAGGCGGAGAGCTCTCT 3480
QY 3481 GACACGCTCAATGACTCTGAGAGAGAGAGATGCTATGCTTGAATGAATGCCGAAGCTTA 3540
DB 3481 GACACGCTCAATGACTCTGAGAGAGAGAGATGCTATGCTTGAATGAATGCCGAAGCTTA 3540
QY 3541 CAGCAGAGCTGAGAGCTGAACGAGAGCTCAACAGAGGCTTCTGAGAGCAGAGCCAA 3600
DB 3541 CAGCAGAGCTGAGAGCTGAACGAGAGCTCAACAGAGGCTTCTGAGAGCAGAGCCAA 3600
QY 3601 TTACAGCAGCAGATGAGACTGTCAGAGAAATGACATTTTCCGTCTGACTCAAGAGCTGCA 3660
DB 3601 TTACAGCAGCAGATGAGACTGTCAGAGAAATGACATTTTCCGTCTGACTCAAGAGCTGCA 3660
QY 3661 GAAGCTCTAGATCGGAGCTGATCTACTGAGACAGAGAGAGAGTGAAGTATGAGCTG 3720
DB 3661 GAAGCTCTAGATCGGAGCTGATCTACTGAGACAGAGAGAGAGTGAAGTATGAGCTG 3720
QY 3721 GAAACCAATTCAGGTTCTCTATTTCTATGAAAGGTAAGTAAGAGGACATTTCTCAA 3780
DB 3721 GAAACCAATTCAGGTTCTCTATTTCTATGAAAGGTAAGTAAGAGGACATTTCTCAA 3780
QY 3781 CAAACCAACTCATTTGATTTTCTGCAAGCCAAATGAGACCAACTGCTTAAAGAGAGAG 3840
DB 3781 CAAACCAACTCATTTGATTTTCTGCAAGCCAAATGAGACCAACTGCTTAAAGAGAGAG 3840
QY 3841 GTTCTCTGCAAGTACATGAGTGAAGTGGCCCTGAGAGAGAGAGAGAGTCTGCTGTCGA 3900
DB 3841 GTTCTCTGCAAGTACATGAGTGAAGTGGCCCTGAGAGAGAGAGAGAGTCTGCTGTCGA 3900
QY 3901 GAGCTAAGAGAGAGGCTTTCAGAGAGCCGACATGAGCTCCGCTCCGCGGAGAGAGAGCT 3960
DB 3901 GAGCTAAGAGAGAGGCTTTCAGAGAGCCGACATGAGCTCCGCTCCGCGGAGAGAGAGCT 3960
QY 3961 GCCCAGCGCAAG 4020
DB 3961 GCCCAGCGCAAG 4020
QY 4021 ATGCGCATGTCGCGCATGTCGAGTGGCCAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4080
DB 4021 ATGCGCATGTCGCGCATGTCGAGTGGCCAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4080
QY 4081 GCCCGGCGATTCAGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4140
DB 4081 GCCCGGCGATTCAGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4140
QY 4141 AAGGAG 4200
DB 4141 AAGGAG 4200
QY 4201 ACAAGT 4260
DB 4201 ACAAGT 4260
QY 4261 GAATGTGAGT 4320
DB 4261 GAATGTGAGT 4320
QY 4321 CCTGCTGAATGT 4380
DB 4321 CCTGCTGAATGT 4380
QY 4381 GGTCTCAGACCAAG 4440
DB 4381 GGTCTCAGACCAAG 4440

DB 4381 GGTCTCAGACCAAG 4440
QY 4441 AGGAATTAACAAAG 4500
DB 4441 AGGAATTAACAAAG 4500
QY 4501 AAAGTCTCATTTATGACATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4560
DB 4501 AAAGTCTCATTTATGACATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4560
QY 4561 GAGCTGAGCTTCCCGAG 4620
DB 4561 GAGCTGAGCTTCCCGAG 4620
QY 4621 GCAAAATACAGCAAG 4680
DB 4621 GCAAAATACAGCAAG 4680
QY 4681 ACCTGCTGAGCCGAG 4740
DB 4681 ACCTGCTGAGCCGAG 4740
QY 4741 TGGGTCAACGCGCTTAAGATGATGTCGAGTGGAGAGAGAGAGAGAGAGAGAGAGAG 4800
DB 4741 TGGGTCAACGCGCTTAAGATGATGTCGAGTGGAGAGAGAGAGAGAGAGAGAGAGAG 4800
QY 4801 GCTGATGCTTAACTGCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4860
DB 4801 GCTGATGCTTAACTGCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4860
QY 4861 ATGAATGACAGCTGAG 4920
DB 4861 ATGAATGACAGCTGAG 4920
QY 4921 TAGGCTGAGATGCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4980
DB 4921 TAGGCTGAGATGCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4980
QY 4981 CAAATTTATTTATTAACAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5040
DB 4981 CAAATTTATTTATTAACAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5040
QY 5041 CTGTGTCTGTGACGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5100
DB 5041 CTGTGTCTGTGACGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5100
QY 5101 CAGCCGACATCTCAACCAATTTTGAAGCTGTCAAGAGAGAGAGAGAGAGAGAGAGAG 5160
DB 5101 CAGCCGACATCTCAACCAATTTTGAAGCTGTCAAGAGAGAGAGAGAGAGAGAGAGAG 5160
QY 5161 GGCAGAGTGAAG 5220
DB 5161 GGCAGAGTGAAG 5220
QY 5221 CGCTACAG 5280
DB 5221 CGCTACAG 5280
QY 5281 TGCAGCTGTATCCATTACCAATTACAGATCTCTATTGGAACCAATTAATTTAGGAA 5340
DB 5281 TGCAGCTGTATCCATTACCAATTACAGATCTCTATTGGAACCAATTAATTTAGGAA 5340
QY 5341 ATGAGACATGAAG 5400
DB 5341 ATGAGACATGAAG 5400
QY 5401 CCTGCTGT 5460
DB 5401 CCTGCTGT 5460
QY 5461 GCAGGAG 5520
DB 5461 GCAGGAG 5520

Db	5461	GCAGGGACCGAGAGAGACTCTTGCTGTGTTTCCACGAATTTGGAGTGTGGTAGATCT	5520
QY	5521	TACGGAAAGCGTAGCCGCAGACAGCCGATCTCAAGTGAAGTCCGTTACTTTGGCCCTTGGCC	5580
Db	5521	TACGGAAAGCGTAGCCGCAGACAGATCTCAAGTGAAGTCCGTTACTTTGGCCCTTGGCC	5580
QY	5581	TACAGAGAACCTTATCTGTTTGTGACCACTTCAACTCACTGGAAGTAATTGAGATCCAG	5640
Db	5581	TACAGAGAACCTTATCTGTTTGTGACCACTTCAACTCACTGGAAGTAATTGAGATCCAG	5640
QY	5641	GCAGCGTCTCAGAGAGGAGACCCCTGGCCGAGCGTACCTGGAATCCGACACCCCGCGTAC	5700
Db	5641	GCAGCGTCTCAGAGAGGAGACCCCTGGCCGAGCGTACCTGGAATCCGACACCCCGCGTAC	5700
QY	5701	CTGGGCCCTGCGCATTTCTCAGAGAGCATTTACTTGGCGTCCCTATACAGAGATTAATTA	5760
Db	5701	CTGGGCCCTGCGCATTTCTCAGAGAGCATTTACTTGGCGTCCCTATACAGAGATTAATTA	5760
QY	5761	AGGCGTATTTGCTGACGAGGAAACCTCGTGAAGAGTCCGGCACTGAAACCAACCGGGGC	5820
Db	5761	AGGCGTATTTGCTGACGAGGAAACCTCGTGAAGAGTCCGGCACTGAAACCAACCGGGGC	5820
QY	5821	CCGTCACACTCCCGCAG 5837	
Db	5821	CCGTCACACTCCCGCAG 5837	

RESULT 15.
ADf60993
ID ADF60993 standard; cDNA; 6162 BP.
XX
XX ADF60993;
XX
DT 12-FEB-2004 (first entry)
XX
DE ORF of pain associated human gene, 2207.
XX
KW Pain modulation; pain disorder; painful disorder; potassium channel;
KW kinase expression; inflammatory pain; chronic pain; neuropathic pain;
KW causalgia; fibromyalgia; cancer pain; migraine; headache; tissue pain;
KW analgesic; antiinflammatory; gene; ss.
XX
OS Homo sapiens.
XX
Key Location/Qualifiers
FT CDS
FT 1..6162
FT /*tag= a
FT /product= "2207 polypeptide"
XX
PN US2003153525-A1.
XX
PD 14-AUG-2003.
XX
PF 19-DEC-2002; 2002US-00325430.
XX
PR 19-DEC-2001; 2001US-0341953P.
XX
PA (MILL-) MILLENNIUM PHARM INC.
XX
PI Silos-Santiago I, Rosenfeld JB;
XX
DR WPI; 2003-897732/82.
XX
XX P-PSDB; ADF60994.
PT
PT Identifying a compound capable of treating a pain disorder comprises
PT assaying the ability of the compound to modulate specific, e.g., kinases
PT or potassium channel, nucleic acid expression or polypeptide activities.
XX
PS Disclosure; SEQ ID NO 11; 80pp; English.
XX
CC The present invention relates to a method for identifying a compound
CC capable of modulating pain or painful disorders. The method comprises
CC assaying the ability of the compound to modulate specific nucleic acid

expression or polypeptide activity e.g. potassium channel, or kinase expression/activity. The method and compounds are useful for treating pain or painful disorders e.g. inflammatory pain, chronic pain, neuropathic pain, causalgia, fibromyalgia, cancer pain, migraine/headache pain and tissue pain. The present sequence encodes a human protein associated with pain.

SQ Sequence 6162 BP; 1746 A; 1528 C; 1680 G; 1208 T; 0 U; 0 Other;

Query Match	91.8%;	Score 5661.4;	DB 10;	Length 6162;
-------------	--------	---------------	--------	--------------

Matches 5944; Conservative 0; Mismatches 16; Indels 243; Gaps 3;

QY 1 ATGTTGAAGTTC AATAATGAGCGCGGAATCCTTGGATGCTGGTCTGCTGAACCATTT 60

Db 1 ATGTTGAGTTC AATATGAGCCCGGAATCCTTGGATGCTGGTCTGCTGAACCAT 60

61 GCCAGCCGGGCTCCAGGCTGAATCTGTCTTCCAGGGAAACCACTTTATGAC TCA 120

b1 GCGAGGCGCTCCAGGCTTATTCGTCTCAGGGGAACCACTTTATTGACTTCAA 120

121 190

[illegible]

100

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

[illegible][illegible]

vy 301 GACACICATBCIATIGTAAAGIGAI GTAGTAGTAGGAGCII IAT IGGCCCAAGAGCGGII ICA 420

501 GACACACACGCTATGAAAGAGGCTTTATGGCCACGAGCAGGTTTCA 420

421 TTTTGGAGAGCGGAACATATATCTCGAAGCACAGCCCTGGATCCCCCAATTA 480

Db 421 TTTTGTGAGGAGACGGACATATTATCTCGAAGCACACAGCCCCGTGGATCCCCCAATTA 480

481 CAGTATGCCCTTTCAGGACAAAAATCACCTTTATCTGGTCATGGAATATCAGCCTGAGGG 540

Db 481 CAGTATGCCCTTTCAGGACAAATCACCCTTTATCTGATGGAGGATATCAGCCTGGAGG 540

541 GACTTGTCTGCACTTTTGAATAGATATGAGGACCACTTAGATGAATAACCTGATACAGTTT 600

Db 541 GACTTGTCTCACTTTTGATAGATATGAGGACCAGTTAGATGA AAAACCTGATACAGTTT 600

601 TACCTAGCTGAGCTGATTTGGCTGTTACACAGCGTTCATCTGATGGGATACGTGCATCGA 660

Db 601 TACCTAGCTGAGCTGATTGCTGTTACACAGCGTTCATCTGATGGGATACGTCATCGA 660

661 GACATCAAGCCTGAGACATTCTGTTGACCGCACAGACACATCAAGCTGTGATTTT 720

Db 661 GACATCAAGCCTGAGACATTCTCGTGAACCGACAGGACACATCAAGCTGGTGAATTT 720

721 GGATCTGCCCGCAAAATGAATTCAACAAGATGGTGAATGCCAAACTCCCGATTGGGACC 780

Db 721 GGATCTGCCCGAATTC AACAAGATGTTCAACAGATGTTGTAATGCCAACTCCGATTGGGACC 780

781 CCAGATTACATGGCTCTGAACTGCTGACATGTGATGAACGGGATGGAAAAGGCACCTAC 840

[illegible][illegible][illegible]

QY 901 TCCCTTCGAGAGGGAACTCTGCGAGAACCTTCAATAATTATGAATTTCCAGCG 960
Db 901 TCCCTTCGAGAGGGAACTCTGCGAGAACCTTCAATAATTATGAATTTCCAGCG 960
QY 961 TTTTGAATTTTCAGATGACCCCAAGTAGCAGTACTTCTTGATCTGAATTCAGAGC 1020
Db 961 TTTTGAATTTTCAGATGACCCCAAGTAGCAGTACTTCTTGATCTGAATTCAGAGC 1020
QY 1021 TTGTTGTGCGGCGCAAGAAAGAGACTGAAGTTTGAAGGTTCTGCTGCATCTTTCTTC 1080
Db 1021 TTGTTGTGCGGCGCAAGAAAGAGACTGAAGTTTGAAGGTTCTGCTGCATCTTTCTTC 1080
QY 1081 TCTPAAATTTGACCTGGAAACAATCTGTAACCTCTCTCCCTCTGTTTCCACCTCAAG 1140
Db 1081 TCTPAAATTTGACCTGGAAACAATCTGTAACCTCTCTCCCTCTGTTTCCACCTCAAG 1140
QY 1141 TCTGACGATGACCTCCAAATTTTGATGAACCAAGAGAAATTTGTTGGTTTCACTCTCT 1200
Db 1141 TCTGACGATGACCTCCAAATTTTGATGAACCAAGAGAAATTTGTTGGTTTCACTCTCT 1200
QY 1201 CCGTGCCAGCTGAGCCCTCAGGCTTCTCGGTTGAAGAACTGCCGTTTGTGGGTTTTCG 1260
Db 1201 CCGTGCCAGCTGAGCCCTCAGGCTTCTCGGTTGAAGAACTGCCGTTTGTGGGTTTTCG 1260
QY 1261 TACAGCAAGGCACTGGGGATTCTTGTAATCTGAATCTGTTGTGCGGTCTGACTCC 1320
Db 1261 TACAGCAAGGCACTGGGGATTCTTGTAATCTGAATCTGTTGTGCGGTCTGACTCC 1320
QY 1321 CCGGCCAAGCTAGCTGCATGGAAGAACTTCTCATCAAAAGCAAGAGCTTCAAGAC 1380
Db 1321 CCGGCCAAGCTAGCTGCATGGAAGAACTTCTCATCAAAAGCAAGAGCTTCAAGAC 1380
QY 1381 TCTCAGAGCAAGTGTCTCAAGATGAGCAAGAAATGACCCGTTTACATCCGAGTGTCA 1440
Db 1381 TCTCAGAGCAAGTGTCTCAAGATGAGCAAGAAATGACCCGTTTACATCCGAGTGTCA 1440
QY 1441 GAGGTGAGGCTGTGCTTAACTCAGAGAGAGGTGAGGTGAAGGCTCTGAGACTCAGAGA 1500
Db 1441 GAGGTGAGGCTGTGCTTAACTCAGAGAGAGGTGAGGTGAAGGCTCTGAGACTCAGAGA 1500
QY 1501 TCCCTCTGAGAGAGCACTTGTCTAATCACTCAAGAAATGCAATGCACTTAAAGCAAGT 1560
Db 1501 TCCCTCTGAGAGAGCACTTGTCTAATCACTCAAGAAATGCAATGCACTTAAAGCAAGT 1560
QY 1561 TTGGAGCAAGCAAGATGAGAGTGTCCAGAGAGATCAAGAGCACTGAGCTTCTCAT 1620
Db 1561 TTGGAGCAAGCAAGATGAGAGTGTCCAGAGAGATCAAGAGCACTGAGCTTCTCAT 1620
QY 1621 GATATCAGAGAGCAAGAGCCGGAAGCTCAAGAAATCAAGAGCAAGAGTCAAGAGCTCAA 1680
Db 1621 GATATCAGAGAGCAAGAGCCGGAAGCTCAAGAAATCAAGAGCAAGAGTCAAGAGCTCAA 1680
QY 1681 GTGGAGAAATGAGGTGATGATGATCAAGTTGGAAGAGATCTTGTCTCAGCAAGAGA 1740
Db 1681 GTGGAGAAATGAGGTGATGATGATCAAGTTGGAAGAGATCTTGTCTCAGCAAGAGA 1740
QY 1741 CCGAGTATCTCTAAGATCTGAGCTGAGAGGTCTGAGCTGCTGCTGCTGAAGAAATTAAG 1800
Db 1741 CCGAGTATCTCTAAGATCTGAGCTGAGAGGTCTGAGCTGCTGCTGCTGAAGAAATTAAG 1800
QY 1801 CCGAAGAGCAAGATGTCAGCATTAACCTGTTGAAGCTTAAGATCAAGAGAAAGCTGAA 1860
Db 1801 CCGAAGAGCAAGATGTCAGCATTAACCTGTTGAAGCTTAAGATCAAGAGAAAGCTGAA 1860
QY 1861 GTGGAGAAATGAGGTGATGATGATCAAGTTGGAAGAGATCTTGTCTCAGCAAGAGA 1920
Db 1861 GTGGAGAAATGAGGTGATGATGATCAAGTTGGAAGAGATCTTGTCTCAGCAAGAGA 1920
QY 1921 CTCGAGAGAAATGAGAGGCTGTAAAGCAAGCAAGAGGCAAGAGCTGCTGAG 1980
Db 1921 CTCGAGAGAAATGAGAGGCTGTAAAGCAAGCAAGAGGCTGCTGAG 1980
QY 1981 AATATCCGCAAGGAGAGAGCGAGGAGGAGGCTGAGAAAGCTGAGAAACGAGAG 2040

Db 1945 -----GAAAAGAGCGAGCCGAGAGGAGCTGGAGAAAGCTGCAAGAACGAGAG 1992
QY 2041 GATTCCTCTGAAGAGCATCAGAAAGAGCTGTGAGCTGAGAGAACCGGCAATTCCTG 2100
Db 1993 GATTCCTCTGAAGAGCATCAGAAAGAGCTGTGAGAGCTGAGAGAACCGGCAATTCCTG 2052
QY 2101 GAGAACAGGTAAAGAGACTGAGAGCAATGAGCGTGAAGAGAAACAGACTGAAGATGAC 2160
Db 2053 GAGAACAGGTAAAGAGACTGAGAGCAATGAGCGTGAAGAGAAACAGACTGAAGATGAC 2112
QY 2161 ATCCAGCAAAATCCCAAGATCCAGAGATGAGCTGATTAATTTCTGAGCTGAGAGAG 2220
Db 2113 ATCCAGCAAAATCCCAAGATCCAGAGATGAGCTGATTAATTTCTGAGCTGAGAGAG 2172
QY 2221 AAAATGAGAGGCGCAAGTCTCAGCCAGCACTGAGAGTGAAGTGAACCTGAAACAGAGAG 2280
Db 2173 AAAATGAGAGGCGCAAGTCTCAGCCAGCACTGAGAGTGAAGTGAACCTGAAACAGAGAG 2232
QY 2281 CAGCACTATGAGAGAAAGATTAAGTGTGACATGATTAAGAAAGACTGCTGAC 2340
Db 2233 CAGCACTATGAGAGAAAGATTAAGTGTGACATGATTAAGAAAGACTGCTGAC 2292
QY 2341 AAGAGCACTGAGAGAACTGATGAGAGCAAGAGAGAGGCGCTGAGAGAGGCAAA 2400
Db 2293 AAGAGCACTGAGAGAACTGATGAGAGCAAGAGAGAGGCGCTGAGAGAGGCAAA 2352
QY 2401 ATTCTCAGCAAGAGAGGATGATCAATGCTATGAGATTCAGAGATCAGATCCCTGAA 2460
Db 2353 ATTCTCAGCAAGAGAGGATGATCAATGCTATGAGATTCAGAGATCAGATCCCTGAA 2412
QY 2461 CAGAGATTTGGAATCTGTCTGAGAGCCAAATTAACCTGCAAGAAATGACAGTCTTTTACC 2520
Db 2413 CAGAGATTTGGAATCTGTCTGAGAGCCAAATTAACCTGCAAGAAATGACAGTCTTTTACC 2472
QY 2521 CAAAGGAACTGAAGGCGCAAGAGAGATGATTTCTGAATCTCAGGCAACGAAATTTTAC 2580
Db 2473 CAAAGGAACTGAAGGCGCAAGAGAGATGATTTCTGAATCTCAGGCAACGAAATTTTAC 2532
QY 2581 CTGAGAGCAAGGCTGGAGAGTTGAGAGGCGCAGAACCGAAATCTGAGAGAGAGCTGAG 2640
Db 2533 CTGAGAGCAAGGCTGGAGAGTTGAGAGGCGCAGAACCGAAATCTGAGAGAGAGCTGAG 2592
QY 2641 AAGATCAGCCACCAAGAGCAAGATGATCAAGATCTGCTGAGAACTGAGAGCAAGATTTG 2700
Db 2593 AAGATCAGCCACCAAGAGCAAGATGATCAAGATCTGCTGAGAACTGAGAGCAAGATTTG 2652
QY 2701 CCGAGGCTCAGTCTGAGAGCAAGAGAGCAAGAACTGAGAGCTCAAGGCGCAGCTCAGAG 2760
Db 2653 CCGAGGCTCAGTCTGAGAGCAAGAGAGCAAGAACTGAGAGCTCAAGGCGCAGCTCAGAG 2712
QY 2761 CTACAGCTCTCCCTGAGAGAGGCGCAGAGTCAAGTTGACACCCCTGAGAGGCTGACAGGCGG 2820
Db 2713 CTACAGCTCTCCCTGAGAGAGGCGCAGAGTCAAGTTGACACCCCTGAGAGGCTGACAGGCGG 2772
QY 2821 GCCCTGAGAGGAGCTTCCGCAAGAGCAAGAGCTGAGAGAGCAAGAGAGCT 2880
Db 2773 GCCCTGAGAGGAGCTTCCGCAAGAGCAAGAGCTGAGAGAGCAAGAGAGCT 2832
QY 2881 GAAAGAGATTCAGAGCACTCAGGCAATGAGATGAAATTCAGGCGCAATTTGATGCT 2940
Db 2833 GAAAGAGATTCAGAGCACTCAGGCAATGAGATGAAATTCAGGCGCAATTTGATGCT 2892
QY 2941 CTTCTGTAACGCTGTACTGTATATCAAGAGCTGAGAGAGAGCTTAACAGCTGACCGAG 3000
Db 2893 CTTCTGTAACGCTGTACTGTATATCAAGAGCTGAGAGAGAGCTTAACAGCTGACCGAG 2952
QY 3001 GACACGCTGAACTCAACCAAACTTACTTGTGCAAGCAATCTGAGTGAAGCTTCT 3060
Db 2953 GACACGCTGAACTCAACCAAACTTACTTGTGCAAGCAATCTGAGTGAAGCTTCT 3012
QY 3061 GAGCGCAACGACGAGATTGTACACTGCAAGTGAAGTGAACATCTCGCGGAGATC 3120

Db 3013 GGGCCAAACGACGATTGTGTAACAACCTGGAGAGTGAAGTGAACATCTCCGCGGGAGATC 3072
QY 3121 ACCGAACGAGAGATGCAGCTTACCGACCAAGAAAGAAACGATGAGGCTCTGAAAGCAACG 3180
Db 3073 ACCGAACGAGAGATGCAGCTTACCGACCAAGAAACGATGAGGCTCTGAAAGCAACG 3132
QY 3181 TGCACATGCTGGAGAAACAGGTGATGGAATTTGAGAGCCCTTAACGATGAGTGCCTGAA 3240
Db 3133 TGCACATGCTGGAGAAACAGGTGATGGAATTTGAGAGCCCTTAACGATGAGTGCCTGAA 3192
QY 3241 AAAGAGCGGAGTGGAGAGGCTGGAGAGGCTCTGGGTGATGAGAAATCCAGTTTGAG 3300
Db 3193 AAAGAGCGGAGTGGAGAGGCTGGAGAGGCTCTGGGTGATGAGAAATCCAGTTTGAG 3252
QY 3301 TGTGGGTTTGAAGACTGTCAGAGAGTCTTGACAACCAAGAAACAGACAGGCGAGAGCC 3360
Db 3253 TGTGGGTTTGAAGACTGTCAGAGAGTCTTGACAACCAAGAAACAGACAGGCGAGAGCC 3312
QY 3361 GATCAGCGGATCAACCGATCTCCGACAGTGTGGAGCTGGCACTGCAAGAGACACAGGCT 3420
Db 3313 GATCAGCGGATCAACCGATCTCCGACAGTGTGGAGCTGGCACTGCAAGAGACACAGGCT 3372
QY 3421 GAGATTCTCGCTCTGACAGAGCTCTCAAGAGCAGAGAGCTGAAGGCGAGAGCTCTCT 3480
Db 3373 GAGATTCTCGCTCTGACAGAGCTCTCAAGAGCAGAGAGCTGAAGGCGAGAGCTCTCT 3432
QY 3481 GACAGCTCAATGACTCTGAGAAAGACATGCTATGCTTGAATGAATGCGCCGAAAGCTTA 3540
Db 3433 GACAGCTCAATGACTCTGAGAAAGACATGCTATGCTTGAATGAATGCGCCGAAAGCTTA 3492
QY 3541 CACGAGAGCTGAGACTGAAACAGAGCTCAACAGAGGCTTTCGAAAGAGCAAGCCAA 3600
Db 3493 CACGAGAGCTGAGACTGAAACAGAGCTCAACAGAGGCTTTCGAAAGAGCAAGCCAA 3552
QY 3601 TTACAGCAGAGATGAGACTGCAAGAAATACATTTTCCGCTGACTCAAGAGCTGCA 3660
Db 3553 TTACAGCAGAGATGAGACTGCAAGAAATACATTTTCCGCTGACTCAAGAGCTGCA 3612
QY 3661 GAACTCTAGATGCGGCTGATCTAAGAGACAGAAAGAAAGTGAAGTATCAAGCTG 3720
Db 3613 GAACTCTAGATGCGGCTGATCTAAGAGACAGAAAGAAAGTGAAGTATCAAGCTG 3672
QY 3721 GAAACATTCAGGTTCTTAATTCATGAAAGAGTGAAGTGAAGGCAATATTTCTCA 3780
Db 3673 GAAACATTCAGGTTCTTAATTCATGAAAGAGTGAAGTGAAGGCAATATTTCTCA 3732
QY 3781 CAAACCAACTCATTTGATTTTCTGAAGCCAAATGGAACCACTGTGTAAGAAAGAAA 3838
Db 3733 CAAACCAACTCATTTGATTTTCTGAAGCCAAATGGAACCACTGTGTAAGAAAGAAA 3792
QY 3839 -----AGTTTCTCTGCAGTAC 3855
Db 3793 GGTATTATTAGTGAACGAAAGAGACCTGCTTACCAACAGGTTCTCTGCAGTAC 3852
QY 3856 AATGAGCTGAAGTGGCCCTGGAGAAAGAGAAAGCTCGCTGTCAAGACTGAGAAAGCC 3915
Db 3853 AATGAGCTGAAGTGGCCCTGGAGAAAGAGAAAGCTCGCTGTCAAGACTGAGAAAGCC 3912
QY 3916 CTTTCAGAAAGCCCGCATGAGCTCGGATCGGCCGGGAGGAAAGCTGCCACCGCAAGCA 3975
Db 3913 CTTTCAGAAAGCCCGCATGAGCTCGGATCGGCCGGGAGGAAAGCTGCCACCGCAAGCA 3972
QY 3976 ACGGACCAACCAACCCATCAAGCCAGCAACCGCAGGAGCAGAGATGCGCATGTCCGC 4035
Db 3973 ACGGACCAACCAACCCATCAAGCCAGCAACCGCAGGAGCAGAGATGCGCATGTCCGC 4032
QY 4036 ATTCGTGGGTCCGCAAGAGCAACCAAGCCAGTGCATGAGCTGTGGCCCGCATCCAGC 4095
Db 4033 ATTCGTGGGTCCGCAAGAGCAACCAAGCCAGTGCATGAGCTGTGGCCCGCATCCAGC 4092
QY 4096 CCGAGAAAGAGTCTTCAACTCCAGAGAAATTTAGTGGGCTTTAAGAAAGCATGAC 4155
Db 4093 CCGAGAAAGAGTCTTCAACTCCAGAGAAATTTAGTGGGCTTTAAGAAAGCATGAC 4152

QY 4156 CACATATTTCTTCAACCGATTCAACGTAGACTGAACATGCGAGCCACAAAGTGTCTGTG 4215
Db 4153 CACATATTTCTTCAACCGATTCAACGTAGACTGAACATGCGAGCCACAAAGTGTCTGTG 4212
QY 4216 TGTGGATACCGTGCACTTTTGAACCCAGGCAATCCAAATGTCCTGGAATGTCAGGTATG 4275
Db 4213 TGTGGATACCGTGCACTTTTGAACCCAGGCAATCCAAATGTCCTGGAATGTCAGGTATG 4272
QY 4276 TGTCAACCCAAAGTCTCAAGTCTTTCACGCAACCTGCGGCTTGCCTGTGATATGCC 4335
Db 4273 TGTCAACCCAAAGTCTCAAGTCTTTCACGCAACCTGCGGCTTGCCTGTGATATGCC 4332
QY 4336 ACACACTTCAACGAGGCTTTCGCGGTGACAAATGAACTTCCAGGTCTCCAGACCAAG 4395
Db 4333 ACACACTTCAACGAGGCTTTCGCGGTGACAAATGAACTTCCAGGTCTCCAGACCAAG 4392
QY 4396 GAGCCCAAGCAGAGCTTTCGACCTGGAAGGTTGATGAAGTGTCCGGAATTAACAAACGA 4455
Db 4393 GAGCCCAAGCAGAGCTTTCGACCTGGAAGGTTGATGAAGTGTCCGGAATTAACAAACGA 4452
QY 4456 GGAACCAAGGCTGGGACAGAAAGTACATTTGCTCTGAGGAGTCAAAAGTCTCATTTAT 4515
Db 4453 GGAACCAAGGCTGGGACAGAAAGTACATTTGCTCTGAGGAGTCAAAAGTCTCATTTAT 4512
QY 4516 GACATGAAGCCAGAGAGCTGGAACAGAGCCGGTGAAGAAATTTGAGCTGTGCTTCCC 4575
Db 4513 GACATGAAGCCAGAGAGCTGGAACAGAGCCGGTGAAGAAATTTGAGCTGTGCTTCCC 4572
QY 4576 GACGGGATGTATCTATTTCAATGCTGCGTGTGCTTCCGAATCCCAATTAAGCCAA 4635
Db 4573 GACGGGATGTATCTATTTCAATGCTGCGTGTGCTTCCGAATCCCAATTAAGCCAA 4632
QY 4636 GCAGATGCTCCATCACTAGTAAGATGAATCAACCCGACACCACTGCTGCGCGGG 4695
Db 4633 GCA----- 4635
QY 4696 AGAACCTTACTTGTAGCTCCAGCTTCTTGACAAACAGCGGTGGTCAACCGCTTGA 4755
Db 4636 ----- 4635
QY 4756 GAATCAGTTGTGCAGGTGGAGAGTTCCTAGGAGAAAGCAGAACTGATGCTAAACTG 4815
Db 4636 -----GAAAAAGCAGAACTGATGCTAAACTG 4662
QY 4816 CTTGAAACTCCCTGTGAAACTGGAAGGTGATGACCTGTAGACATGAACTGACGCTG 4875
Db 4663 CTTGAAACTCCCTGTGAAACTGGAAGGTGATGACCTGTAGACATGAACTGACGCTG 4722
QY 4876 CCGTTCAGTGACCAAGTGGTGTGGTGGGACCCGAGAAAGGCTCTAGCCCTGAATGTG 4935
Db 4723 CCGTTCAGTGACCAAGTGGTGTGGTGGGACCCGAGAAAGGCTCTAGCCCTGAATGTG 4782
QY 4936 TTGAAAAAATCCCTTAACCATGTCCCAAGAAATTTGAGAGCTTCCAAATTTATATATAC 4995
Db 4783 TTGAAAAAATCCCTTAACCATGTCCCAAGAAATTTGAGAGCTTCCAAATTTATATATAC 4842
QY 4996 AAGAACCTGGAAGAGTACTCATGATAGCAGAGAAAGACGGGCACTGTGCTTGTGAC 5055
Db 4843 AAGAACCTGGAAGAGTACTCATGATAGCAGAGAAAGACGGGCACTGTGCTTGTGAC 4902
QY 5056 GTGAAGAAAGTGAACAGTCCCTGGCCAGTCCCACTTGCCGACGCCAGCATCTGA 5115
Db 4903 GTGAAGAAAGTGAACAGTCCCTGGCCAGTCCCACTTGCCGACGCCAGCATCTGA 4962
QY 5116 CCCAACATTTTGAAGCTGTCAAGGGCTGCACTTGTTTGGGCAAGCAAGTTGAGAAC 5175
Db 4963 CCCAACATTTTGAAGCTGTCAAGGGCTGCACTTGTTTGGGCAAGCAAGTTGAGAAC 5022
QY 5176 GGGCTCTGCATCTGTGAGCCATGCCCAGCAAAAGTGTGATCTCTCGCTTAACAGAAAC 5235
Db 5023 GGGCTCTGCATCTGTGAGCCATGCCCAGCAAAAGTGTGATCTCTCGCTTAACAGAAAC 5082

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 11, 2005, 16:31:18 ; Search time 256 Seconds

(without alignment)
4108.635 Million cell updates/sec

Title: US-10-791-666-2

Sequence: 1 MLKFKYGRAPLDAGAAEPL.....AGAVRPLSGVKNKWDQSSV 2054

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt_03.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Length	ID	Description
1	10230.5	97.5	2027	086U09	086u09 homo sapien
2	10109.5	96.3	2055	088938	088938 mus musculu
3	8161	77.8	1641	088528	088528 mus musculu
4	7935	75.6	1597	CTRO_MOUSE	P49025 mus musculu
5	7851.5	74.8	1618	090X19	090x19 rattus norv
6	6575	62.7	1286	CTRO_HUMAN	014578 homo sapien
7	5057.5	48.2	1032	08C1D1	08c1d1 mus musculu
8	2492	23.7	482	06XU08	06xuh8 homo sapien
9	2175.5	20.7	494	088937	088937 mus musculu
10	2160.5	20.6	448	088527	088527 rattus norv
11	1822	17.4	1854	09VU78	09vct8 drosophila
12	1524.5	14.5	1732	054874	054874 rattus norv
13	1489	14.2	1719	096XK3	096xk3 homo sapien
14	1440	13.7	1760	09U0U5	09u0u5 homo sapien
15	1438	13.7	1711	09Y5S2	09y5s2 homo sapien
16	1420.5	13.5	1733	07TT49	07t492 rattus norv
17	1419.5	13.5	1702	054875	054875 rattus norv
18	1418.5	13.5	1713	07IT50	07it50 mus musculu
19	1413.5	13.5	1638	086XK2	086xk2 homo sapien
20	1409.5	13.4	1638	08IMQ7	08imq7 homo sapien
21	1401	13.4	1573	07PVB7	07pvt7 anopheles g
22	1340.5	12.8	1551	06DT37	06dt37 homo sapien
23	1316.5	12.5	1582	001583	001583 caenorhabdi
24	1279	12.2	1369	ROCI_RAT	063644 rattus norv
25	1275.5	12.2	1354	ROCI_MOUSE	070335 mus musculu
26	1270.5	12.1	1637	09W1B0	09w1b0 drosophila
27	1266.5	12.1	1613	044368	044368 drosophila
28	1263.5	12.0	1388	ROC2_BOVIN	028021 bos taurus
29	1260	12.0	1354	ROCI_HUMAN	013464 homo sapien
30	1257	12.0	1388	ROCI_HUMAN	075116 homo sapien
31	1255.5	12.0	1375	090Y37	090y37 brachydanio

32	1250.5	11.9	1388	1	ROC2_MOUSE	P70336 mus musculu
33	1249	11.9	1354	1	ROCI_RABIT	077819 o rho-asso
34	1248.5	11.9	1370	2	073732	073732 xenopus lae
35	1232.5	11.7	1379	1	ROC2_RAT	062868 rattus norv
36	1168	11.1	1390	2	09U779	09u779 drosophila
37	1168	11.1	1390	2	09VXE3	09vxe3 drosophila
38	1167	11.1	1390	2	086UD4	086ud4 drosophila
39	1143	10.9	933	2	086TU1	086tj1 homo sapien
40	1126.5	10.7	1173	2	P92199	P92199 caenorhabdi
41	1075	10.2	865	2	098SN6	098sn6 gallus galli
42	1072	10.2	1386	2	07PRV2	07prv2 anopheles g
43	1060.5	10.1	717	2	08AVM0	08avm0 xenopus lae
44	1003.5	9.6	1231	2	07QIK8	07qik8 anopheles g
45	925	8.8	492	2	086XZ8	086xz8 homo sapien

ALIGNMENTS

RESULT 1
ID 086U09 PRELIMINARY; PRT; 2027 AA.
AC 086U09;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Rho/rac-interacting citron kinase.
GN Name=CIT;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Huang C.Q., Wu S.L., Shan Y.X., Liu S., Xiao P.J.;
RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Belongs to the Ser/Thr protein kinase family.
CC -1- SIMILARITY: Contains 1 PH domain.
DR EMBL; AY257469; APL3528.1; -.
DR HSSP; P31751; IMRV.
DR GO; GO:0005622; C:intracellular; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004674; F:Protein serine/threonine kinase activity; IEA.
DR GO; GO:0005083; F:small GTPase regulatory/interacting protein. .; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0007242; P:intracellular signaling cascade; IEA.
DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR InterPro; IPR001180; Citron.
DR InterPro; IPR005479; Cphp_synch_L_D2.
DR InterPro; IPR002219; DAG_PE_bind.
DR InterPro; IPR011009; Kinase_like.
DR InterPro; IPR001849; PH.
DR InterPro; IPR000961; Pkinase_C.
DR InterPro; IPR000719; Prot_Kinase.
DR InterPro; IPR000861; REM_repeat.
DR InterPro; IPR002290; Ser_thr_kinase.
DR InterPro; IPR008271; Ser_thr_kin_AS.
DR Pfam; PF00130; C1.1; 1.
DR Pfam; PF00780; CNH; 1.
DR Pfam; PF00169; PH; 1.
DR Pfam; PF00069; Pkinase; 1.
DR Pfam; PF00433; Pkinase_C; 1.
DR Prodom; PD000001; Prot_kinase; 1.
DR SMART; SM00109; C1.1.
DR SMART; SM00036; CNH; 1.
DR SMART; SM00233; PH; 1.
DR SMART; SM00220; S_TKc; 1.
DR SMART; SM00133; S_TKc_X; 1.
DR PROSITE; PS00867; CPKase_2; UNKNOWN 1.
DR PROSITE; PS00479; DAG_PE_BIND_DOM_1; 1.
DR PROSITE; PS00081; DAG_PE_BIND_DOM_2; 1.
DR PROSITE; PS00003; PH_DOMAIN; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.

DR PROSITE; PSS0011; PROTEIN KINASE, DOM; 1.
DR PROSITE; PSS0108; PROTEIN KINASE_ST; 1.
KW ATP-binding; Kinase; Serine/Threonine-Protein kinase; Transferase.
SQ SEQUENCE 2027 AA; 231429 MW; 6B1D8C3F661F357B CRC64;

Query Match 97.5%; Score 10230.5; DB 2; Length 2027;

Best Local Similarity 97.1%; Pred. No. 2e-285; Matches 2010; Conservative 1; Mismatches 1; Indels 57; Gaps 2;

QY 1 MLKFKYGARPLDAGAAPIASRASRLNLFQCKPPTQOQMSPLSREGILDALFVLFE 60
DB 1 MLKFKYGARPLDAGAAPIASRASRLNLFQCKPPTQOQMSPLSREGILDALFVLFE 60
QY 61 ECGQPALMTKHYSNFTRKTSDDTIAELQELQPSAKDPVNSLYCCGHPAAYQVYREKATG 120
DB 61 ECGQPALMTKHYSNFTRKTSDDTIAELQELQPSAKDPVNSLYCCGHPAAYQVYREKATG 120
QY 121 DIYAMKVMKKKALLAQOVSFPFEERNILSRSTSPMIPOLQYAFQDNHLYLWVEYQPG 180
DB 121 DIYAMKVMKKKALLAQOVSFPFEERNILSRSTSPMIPOLQYAFQDNHLYLWVEYQPG 180
QY 181 DLISLNRVEDQDENLIQFYLAELILAVSHVLMGVHRDIKENILVDRTGHIKLVD 240
DB 181 DLISLNRVEDQDENLIQFYLAELILAVSHVLMGVHRDIKENILVDRTGHIKLVD 240
QY 241 GSAAKNSNKNVNAKLPITGPDYAPVYLVNMGDGKTTYGLDCDWMSVGVLAYEMTYGR 300
DB 241 GSAAKNSNKNVNAKLPITGPDYAPVYLVNMGDGKTTYGLDCDWMSVGVLAYEMTYGR 300
QY 301 SPAPEGSARTFNINMFQRFKPPDPKYSDFDLIQSLGCGKRLKFBGLCGHPF 360
DB 301 SPAPEGSARTFNINMFQRFKPPDPKYSDFDLIQSLGCGKRLKFBGLCGHPF 360
QY 361 SKIDMNNIRNSPPFPYTLKSDDDTSNFDPEKNSWVSSPCQSPSGFSGBELPFVGS 420
DB 361 SKIDMNNIRNSPPFPYTLKSDDDTSNFDPEKNSWVSSPCQSPSGFSGBELPFVGS 420
QY 421 YSKALGILGRSESVSGLDSPAKTSSMEKLLIKSKELQDSQDKCHMEQEMRLHRRVS 480
DB 421 YSKALGILGRSESVSGLDSPAKTSSMEKLLIKSKELQDSQDKCHMEQEMRLHRRVS 480
QY 481 EYVAVLSQKEVELKASTQSRSLLEODLATYITTESSIKRSLQAPMVSQEDDKALQLH 540
DB 481 EYVAVLSQKEVELKASTQSRSLLEODLATYITTESSIKRSLQAPMVSQEDDKALQLH 540
QY 541 DIBEQSKLOEIKEOEYQAOVEEMRLMMNQLLEBDLVARRRSDLYESELRESLLAAEFK 600
DB 541 DIBEQSKLOEIKEOEYQAOVEEMRLMMNQLLEBDLVARRRSDLYESELRESLLAAEFK 600
QY 601 RKATECQHKLLKAKDQKPEVGEYAKLEKINAEOQLIOELOEKLKAVASTEATELLQ 660
DB 601 RKATECQHKLLKAKDQKPEVGEYAKLEKINAEOQLIOELOEKLKAVASTEATELLQ 660
QY 661 NITQAKERARRELKQONREDSSEGTFRKLVAEERHSLLENKYKRLTETMERENRLKD 720
DB 661 NITQAKERARRELKQONREDSSEGTFRKLVAEERHSLLENKYKRLTETMERENRLKD 720
QY 721 IQTKSQOIQOMADKILELSEKHEEAQVSAOHLLEVHLKQEQHYEEKIKVLNDQIKQDLAD 780
DB 721 IQTKSQOIQOMADKILELSEKHEEAQVSAOHLLEVHLKQEQHYEEKIKVLNDQIKQDLAD 780
QY 781 KETLENNMQHBEBAHEKGIKILSEQKAMINAMDSKIRSLERIVELSEANKLAANSLSFT 840
DB 739 KETLENNMQHBEBAHEKGIKILSEQKAMINAMDSKIRSLERIVELSEANKLAANSLSFT 798
QY 841 QRMNKAQEBMISLROOKFYLETOAGLLEAONKRLLEQOLKISIQHDSDNKRLILELTRL 900
DB 799 QRMNKAQEBMISLROOKFYLETOAGLLEAONKRLLEQOLKISIQHDSDNKRLILELTRL 858
QY 901 REVSLSEHEBQKLEKROLTELQSLQERESQLTALQARALAESQLEQATLELETTAA 960
DB 859 REVSLSEHEBQKLEKROLTELQSLQERESQLTALQARALAESQLEQATLELETTAA 918

QY 961 EERIOALTAHREIQKRPALNSCTVITDLSEQLNLIEDNAELNNONPYLSKOLDEAS 1020
DB 919 EERIOALTAHREIQKRPALNSCTVITDLSEQLNLIEDNAELNNONPYLSKOLDEAS 978
QY 1021 GANDEIVOLREVDHARREITEREMQTSQOKQMEALKTTCMLLEQVNDLEALNDELLE 1080
DB 979 GANDEIVOLREVDHARREITEREMQTSQOKQMEALKTTCMLLEQVNDLEALNDELLE 1038
QY 1081 KERQWEAMNSVIGDEKSGQECRVREIQRMILDEKSGRADQRTITSROYVELAYENKHA 1140
DB 1039 KERQWEAMNSVIGDEKSGQECRVREIQRMILDEKSGRADQRTITSROYVELAYENKHA 1098
QY 1141 EITALLQALKEOKLKAEISLSDKLNDEKKNAMLENNARSLOOGLTEREELKORLLEQAK 1200
DB 1099 EITALLQALKEOKLKAEISLSDKLNDEKKNAMLENNARSLOOGLTEREELKORLLEQAK 1158
QY 1201 LQQQMDLQKNHIFRLTQGLQEALRADLILKTERSDLEYQLENTQVLYSHEKVMEGTISQ 1260
DB 1159 LQQQMDLQKNHIFRLTQGLQEALRADLILKTERSDLEYQLENTQVLYSHEKVMEGTISQ 1218
QY 1261 QTKLIDFLQAKMDQPAKKKK-----VPLQYNELKALEKKAACALEEA 1305
DB 1219 QTKLIDFLQAKMDQPAKKKKGLFSRREKDPALPTQVPLQYNELKALEKKAACALEEA 1278
QY 1306 LQKTRIELSAREEAHRAKTDHPHPSTPATARQOIFMSAIVRSPHOFSAMSLAPSS 1365
DB 1279 LQKTRIELSAREEAHRAKTDHPHPSTPATARQOIFMSAIVRSPHOFSAMSLAPSS 1338
QY 1366 RRKESSTPEEFSRLKERMHNNIPIHFVNLGNMRAKCAVCLDTVHFGQASKCLEQVM 1425
DB 1339 RRKESSTPEEFSRLKERMHNNIPIHFVNLGNMRAKCAVCLDTVHFGQASKCLEQVM 1398
QY 1426 CHEKCSCTCLPATCGLPAYATHTFAFCRDKNNSPELQTEPSSSLHLEGMKVPPNNK 1485
DB 1399 CHEKCSCTCLPATCGLPAYATHTFAFCRDKNNSPELQTEPSSSLHLEGMKVPPNNK 1458
QY 1486 GQGWMDRKTYLVEGSKVLYLNDNAREAGORPVEEPELCIPDGVVSJHGAVASLNTAK 1545
DB 1459 GQGWMDRKTYLVEGSKVLYLNDNAREAGORPVEEPELCIPDGVVSJHGAVASLNTAK 1518
QY 1546 ADVPYTLKMSHPHTCWPGRITLYLAPSPFDKQRTWTALESVAVAGRVSRKAEADAKL 1605
DB 1519 ADVPYTLKMSHPHTCWPGRITLYLAPSPFDKQRTWTALESVAVAGRVSRKAEADAKL 1578
QY 1606 LQNSLLKLBGDRLDNNCTLPSSDQVYLVGTBEGLYALNVLKNSLTHVPGIGAVFOIYII 1665
DB 1579 LQNSLLKLBGDRLDNNCTLPSSDQVYLVGTBEGLYALNVLKNSLTHVPGIGAVFOIYII 1638
QY 1666 KOLEKILMAGEBRALCLVDYKVKQSLAOSHLPAQPDISPMIFEAVKGCCHLFGAKIEN 1725
DB 1639 KOLEKILMAGEBRALCLVDYKVKQSLAOSHLPAQPDISPMIFEAVKGCCHLFGAKIEN 1698
QY 1726 GLCICGAMPSKVYIARYNENLSKYCIKKEIETSEPCSHFTNYSILIGTNKRYEIDMQ 1785
DB 1699 GLCICGAMPSKVYIARYNENLSKYCIKKEIETSEPCSHFTNYSILIGTNKRYEIDMQ 1758
QY 1786 YTLLEFPLDKNDHSLADAVFAASSNSFPVSIYOVNSAGQREBYLLCFHEFGVFDYSYGRS 1845
DB 1759 YTLLEFPLDKNDHSLADAVFAASSNSFPVSIYOVNSAGQREBYLLCFHEFGVFDYSYGRS 1818
QY 1846 RTDDLKWSRLPLAFAREBYLPTVHNSLLEVEIQARSSAGPARAYLIDIPNRYLGPPI 1905
DB 1819 RTDDLKWSRLPLAFAREBYLPTVHNSLLEVEIQARSSAGPARAYLIDIPNRYLGPPI 1878
QY 1906 SSGAIIYLAASYODKLVICCKGNLYVSGTEHHRGSTSRSSPNKSGPTVMEHITKRYA 1965
DB 1879 SSGAIIYLAASYODKLVICCKGNLYVSGTEHHRGSTSRSSPNKSGPTVMEHITKRYA 1938
QY 1966 SSPAPPEGSPHREPTPHRYREGRTELRRDSGGRPLEREKSPGRILSTRERSPARLF 2025
DB 1939 SSPAPPEGSPHREPTPHRYREGRTELRRDSGGRPLEREKSPGRILSTRERSPARLF 1998
QY 2026 BBSRGRLPAAGAVRTPLSOVNVKWDQSSV 2054

Db 1999 EDSRGLPACAVTPTLSQVKNKWDQSSV 2027

RESULT 2

088938 PRELIMINARY; PRT: 2055 AA.
AC 088938; (TRENBLrel. 08, Created)
DT 01-NOV-1998 (TRENBLrel. 08, Last sequence update)
DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)
DE Rho/rac-interacting citron kinase.
GN Name=Cit; Synonyms=Citk;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxId=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Spleen;
RX MEDLINE=99009084; PubMed=9792683; DOI=10.1074/jbc.273.45.29706;
RA Di Cunto F., Calautti E., Hsiao J., Ong L., Topley G., Turco E.,
RA Decto G.P.;
RT "Citron rho-interacting kinase, a novel tissue-specific ser/thr kinase
RT encompassing the Rho-Rac-binding protein Citron.";
RL J. Biol. Chem. 273:29706-29711 (1998).
CC -1- SIMILARITY: Belongs to the Ser/thr protein kinase family.
CC -1- SIMILARITY: Contains 1 PH domain.
DR EMBL: AF068824; AAC72823.1; -.
DR HSSP: P31751; 1MRV.
DR MGD; MGI:105313; Cit.
DR GO; GO:0005622; C:intracellular; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR GO; GO:0005083; F:small GTPase regulatory/interacting protein. . .; IEA.
DR GO; GO:0007242; P:intracellular signaling cascade; IEA.
DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR InterPro: IPR001180; Citron.
DR InterPro: IPR002219; Cphr synth_L_D2.
DR InterPro: IPR002219; DAG_Pe-bind.
DR InterPro: IPR011009; kinase_like.
DR InterPro: IPR011049; PH.
DR InterPro: IPR000961; kinase_C.
DR InterPro: IPR000719; Prot kinase.
DR InterPro: IPR000861; REM Repeat.
DR InterPro: IPR002290; Ser_thr_kinase.
DR InterPro: IPR008271; Ser_thr_kin_AS.
DR InterPro: IPR000533; Tyrosinase.
DR Pfam; PF00130; C1.1; 1.
DR Pfam; PF00780; CNH; 1.
DR Pfam; PF00169; PH; 1.
DR Pfam; PF00069; Kinase; 1.
DR PRINTS; PR00194; TROPOMYOSIN.
DR ProDom; PD000001; Prot kinase; 1.
DR SMART; SM00109; C1.1; 1.
DR SMART; SM00036; CNH; 1.
DR SMART; SM00233; PH; 1.
DR SMART; SM00220; S_TKc; 1.
DR SMART; SM00133; S_TK_X; 1.
DR PROSITE; PS00867; CPSASE_2; UNKNOWN_1.
DR PROSITE; PS00479; DAG_Pe_BIND_DOM_1; 1.
DR PROSITE; PS50081; DAG_Pe_BIND_DOM_2; 1.
DR PROSITE; PS50003; PH_DOMAIN; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 2055 AA; 235480 MW; 2120CB5B454DA940 CRC64;

Query Match 96.3%; Score 10109.5; DB 2; Length 2055;
Best Local Similarity 96.0%; Pred. No. 5.9e-282;

Matches 1974; Conservative 40; Mismatches 39; Indels 3; Gaps 2;

QY 1 MLKFKYGARNPLDAGAAEPIASRASRLNLPFGKRPMTQOQMSPLSEGLDALLFVLF 60
Db 1 MLKFKYGVNPPPEASASEPIASRASRLNLPFGKRPMTQOQMSPLSEGLDALLFVLF 60
QY 61 ECSQPALMKIKHNSFPRKXSDTIAEIQELQPSAKDEPVNSLVCCGFAEIQVYREKATG 120
Db 61 ECSQPALMKIKHNSFPRKXSDTIAEIQELQPSAKDEPVNSLVCCGFAEIQVYREKATG 120
QY 121 DIYMKVKKKALLAQQVSPFEERNIILSTSPWIPOLQYAPQDNKHLVMEYQGG 180
Db 121 DIYMKVKKKALLAQQVSPFEERNIILSTSPWIPOLQYAPQDNKHLVMEYQGG 180
QY 181 DLSLNRYEDQDENIQFYLAELILAVSHVLMGVNHRDKPENTLVDRTGHIKLVDF 240
Db 181 DLSLNRYEDQDENIQFYLAELILAVSHVLMGVNHRDKPENTLVDRTGHIKLVDF 240
QY 241 GSAAKNSNKNVNAKLPIGTPDYAPPEVLTVANGDGKTYGLDCDMSVGYIAYEMTYGR 300
Db 241 GSAAKNSNKN-VDPAKLPIGTPDYAPPEVLTVNMDRGRGTGLDCDMSVGVAYEMTYGK 299
QY 301 SPFAEGTSARTFNNINMFORFLKPPDDPKYSDELDIQSILGQKXRLPBGICHPF 360
Db 300 TPFTGTSARTFNNINMFORFLKPPDDPKYSDELDIQSILGQKXRLPBGICHPF 359
QY 361 SKIDMNNIRNSPPPEVFTLSDDDTSNFDEPKNSWSSPQCQSPGSGEELPFYGF 420
Db 360 ARTDMNNIRNSPPPEVFTLSDDDTSNFDEPKNSWFFILCPAEPPLAFSGEELPFYGF 419
QY 421 YSKALGILGSESVSGLSDPATSSMEKKLILKSKELQSDQCHMEQEMTLHRVS 480
Db 420 YSKALGILGSESVSVSLSDSPAKVSMEKKLILKSKELQSDQCHMEQEMTLHRVS 479
QY 481 EVNAVISOKEVELKASTQSLIBODLATYITCSSIKRLEQARMVSGEDDPAQLLH 540
Db 480 EVNAVISOKEVELKASTQSLIBODLATYITCSSIKRLEQARMVSGEDDPAQLLH 539
QY 541 DIREQSRKQOEIKQEOEYQAVEEMRLMNOLEEDLVASARRSDLYESELRESRLAAEEFK 600
Db 540 DIREQSRKQOEIKQEOEYQAVEEMRLMNOLEEDLVASARRSDLYESELRESRLAAEEFK 599
QY 601 RKATECOHKLKAKDQKPEVGEYAKLEKINAEQQLIQELQEXLEKAVASTATELLO 660
Db 600 RKANECQHKLMKAKDQKPEVGEYSKLEKINAEQQLIQELQEXLEKAVASTATELLO 659
QY 661 NIQAQKRRARELEKLNREDSSGIRKKLVABERHSHLENKYKULETHERENRLKDD 720
Db 660 NIQAQKRRARELEKLNREDSSGIRKKLVABERHSHLENKYKULETHERENRLKDD 719
QY 721 IOTKSOQIQOMADKILEEKEHREBAQVSAOHLFVHLKQKQYBEKIKVLDNOIKKDLAD 780
Db 720 IOTKSOQIQOMADKILEEKEHREBAQVSAOHLFVHLKQKQYBEKIKVLDNOIKKDLAD 779
QY 781 KETLENNMQHBEBAHEKGIKLEQKAMINAMDSKIRSLRQIVELSEANKLAANSLSFT 840
Db 780 KETLENNMQHBEBAHEKGIKLEQKAMINAMDSKIRSLRQIVELSEANKLAANSLSFT 839
QY 841 QRNMKAQOEKISLROOKFLFTQAGLEAQNKLLEQLEKISHODSDKORLLELTR 900
Db 840 QRNMKAQOEKISLROOKFLFTQAGLEAQNKLLEQLEKISHODSDKORLLELTR 899
QY 901 REVSLEHEBOEKLKQLTQLSTLQRESQTLALQAAARALASQALQATELEETABA 960
Db 900 REVSLEHEBOEKLKQLTQLSTLQRESQTLALQAAARALASQALQATELEETABA 959
QY 961 EEBIQALTARDEIQKFDALRNSCTVITDLEBQNLQTEDNAELNNQNFYLSKQJDEAS 1020
Db 960 EEBIQALTARDEIQKFDALRNSCTVITDLEBQNLQTEDNAELNNQNFYLSKQJDEAS 1019
QY 1021 GANDEIVQARSVDHLRREITEREMQITSQKQMEALKTTCTMLEBOVMDLBNDELLE 1080
Db 1020 GANDEIVQARSVDHLRREITEREMQITSQKQMEALKTTCTMLEBOVMDLBNDELLE 1079

QY 1081 KEROMEAMRSVLGDEKSGQFECRVAEILQRMLDTEKOSRABADQRTTESQVVELAVKXENKA 1140
 DB 1080 KEROMEAMRSVLGDEKSGQFECRVAEILQRMLDTEKOSRABADQRTTESQVVELAVKXENKA 1139
 QY 1141 EIALAOLALKEQKLSKASLSDKNDLEKXAMLEMMNARSLOQKLETERELKORLLEBQAK 1200
 DB 1140 EIALAOLALKEQKLSKASLSDKNDLEKXAMLEMMNARSLOQKLETERELKORLLEBQAK 1199
 QY 1201 LQOQMDLQKNIIFELTQGLQBALRADLILKTERSDLEYQLENTIOVLXSHKVKMEGTISQ 1260
 DB 1200 LQOQMDLQKNIIFELTQGLQBALRADLILKTERSDLEYQLENTIOVLXSHKVKMEGTISQ 1259
 QY 1261 QTKLIDFLQAKMDQPAKXKVPLOYNELKALEKXAKCALEBALQKTRIELRSAAEEA 1320
 DB 1260 QTKLIDFLQAKMDQPAKXKVPLOYNELKALEKXAKCALEBALQKTRIELRSAAEEA 1319
 QY 1321 AHRKATDHPHSPATATKROQIAMSALVRSPEHOSANSLAPSPSRKESSTPEEFRRLL 1380
 DB 1320 AHRKATDHPHSPATATKROQIAMSALVRSPEHOSANSLAPSPSRKESSTPEEFRRLL 1379
 QY 1381 KERMEHNIIPRFVNGLMNRATKCAVCDTVHFGROAKSCLECOVMCHPKCSTCLPATCGL 1440
 DB 1380 KERMEHNIIPRFVNGLMNRATKCAVCDTVHFGROAKSCLECOVMCHPKCSTCLPATCGL 1439
 QY 1441 PAEYATHTEAFCDKXNNSPGLQTEBSSSLIEGMKVPNNKRGQOGMDRKIYILEGS 1500
 DB 1440 PAEYATHTEAFCDKXNNSPGLQTEBSSSLIEGMKVPNNKRGQOGMDRKIYILEGS 1499
 QY 1501 KULIYNDEARAGRPVEEFELCLPDQDVSIHGAVGSELANTKADVPYILKXESHPT 1560
 DB 1500 KULIYNDEARAGRPVEEFELCLPDQDVSIHGAVGSELANTKADVPYILKXESHPT 1559
 QY 1561 TCWPGRTLYLAPSPDKOMWTLAESVAVGRRSRKAKADAKLGNLSLKLLEGDRLD 1620
 DB 1560 TCWPGRTLYLAPSPDKOMWTLAESVAVGRRSRKAKADAKLGNLSLKLLEGDRLD 1619
 QY 1621 MNCTLPSPDOVILVGTSEGLYALNVLNKSILTHVIGIGAVFOIYIILKLEKILMAGEERA 1680
 DB 1620 MNCTLPSPDOVILVGTSEGLYALNVLNKSILTHVIGIGAVFOIYIILKLEKILMAGEERA 1679
 QY 1681 LCLVDVKKVQKSLAQSHLPAQPDSPNIFEAIVKCHLFGAKIENGSLCICAMPKSVIIL 1740
 DB 1680 LCLVDVKKVQKSLAQSHLPAQPDSPNIFEAIVKCHLFGAKIENGSLCICAMPKSVIIL 1739
 QY 1741 RYMNENSKYCIKRIETSEPCSCCHFTNYSILIGTNKFYEIDMKQYTLDEFLDKNDHSLA 1800
 DB 1740 RYMNENSKYCIKRIETSEPCSCCHFTNYSILIGTNKFYEIDMKQYTLDEFLDKNDHSLA 1799
 QY 1801 PAVFAASNSFPVSIYVONSAGOREEYILCFHEFGVFDVSGRRSRTDDLKMSRLPLAFA 1860
 DB 1800 PAVFAASNSFPVSIYVONSAGOREEYILCFHEFGVFDVSGRRSRTDDLKMSRLPLAFA 1859
 QY 1861 YREBYLFVTHFNLSLEYEIQARSSAGTPARAYLIDIPPRYLGPALISSGAYILASSYDYL 1920
 DB 1860 YREBYLFVTHFNLSLEYEIQARSSAGTPARAYLIDIPPRYLGPALISSGAYILASSYDYL 1919
 QY 1921 RVLCCCKNLVKESTGTEHHRGSTRSSPNKRGPTTYMEHTTKRYASSPAPPEGSHREP 1980
 DB 1920 RVLCCCKNLVKESTGTEHHRGSTRSSPNKRGPTTYMEHTTKRYASSPAPPEGSHREP 1979
 QY 1981 STPHRY--REGRTIELRRDKSPGRPLEREKSPGRILSTRRERSPALLPEDSSRGLPLAGAV 2038
 DB 1980 STPHRY--REGRTIELRRDKSPGRPLEREKSPGRILSTRRERSPALLPEDSSRGLPLAGAV 2039
 QY 2039 RTPLSQVNAKWDQSSV 2054
 DB 2040 RTPLSQVNAKWDQSSV 2055

RESULT 3
 088528 PRELIMINARY; PRT; 1641 AA.

AC 088528;
 DT 01-NOV-1998 (TReMBLrel. 08, Created)
 DT 01-NOV-1998 (TReMBLrel. 08, last sequence update)
 DT 01-MAR-2004 (TReMBLrel. 26, last annotation update)
 DE CItion-X kinase (Fragment).
 GN Name=Cit.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxId=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RC MEDLINE=96128238; PubMed=8543060; DOI=10.1016/0014-5793(95)01351-2;
 RA Madanle P., Furuyaehiki T., Reid T., Ishizaki T., Watanabe G.,
 RA Morii N., Narumiya S.;
 RT "A novel partner for the GTP-bound forms of rho and rac.";
 RL FBS Lett. 377:243-248(1995).
 CC -1- SIMILARITY: Contains 1 PH domain.
 DR EMBL; AF070066; AAC27933.1; -.
 DR MGI; MGI:105313; Cit.
 DR GO; GO:0005622; C:intracellular; IEA.
 DR GO; GO:0005524; F:ATP binding; IEA.
 DR GO; GO:0016501; F:kinase activity; IEA.
 DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
 DR GO; GO:0005083; F:small GTPase regulatory/interacting protein. .; IEA.
 DR GO; GO:0007242; P:intracellular signaling cascade; IEA.
 DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
 DR InterPro; IPR001180; CItion.
 DR InterPro; IPR005479; Gpnd_synch_L_D2.
 DR InterPro; IPR002219; DAG_PE-bind.
 DR InterPro; IPR001849; PH.
 DR InterPro; IPR000961; Pkinase_C.
 DR InterPro; IPR000861; REM_repeat.
 DR Pfam; PF00130; C1_1; 1.
 DR Pfam; PF00780; CNH; 1.
 DR Pfam; PF00169; PH; 1.
 DR Pfam; PF00433; Pkinase_C; 1.
 DR SMART; SM00109; C1; 1.
 DR SMART; SM00036; CNH; 1.
 DR SMART; SM00233; PH; 1.
 DR SMART; SM00133; S_TK_X; 1.
 DR PROSITE; PS00867; CFSASE_2; UNKNOWN_1.
 DR PROSITE; PS00479; DAG_PE_BIND_DOM_1; 1.
 DR PROSITE; PS50081; DAG_PE_BIND_DOM_2; 1.
 DR PROSITE; PS50003; PH_DOMAIN; 1.
 KW Kinase.
 FT NON_TER
 SQ SEQUENCE 1641 AA; 187658 MW; 6723B20BCA3F22AB CRC64;

Query Match 77.8%; Score 8161; DB 2; Length 1641;
 Best Local Similarity 95.7%; Pred. No. 3e-226;
 Matches 1610; Conservative 15; Mismatches 14; Indels 44; Gaps 2;
 QY 374 PFVPTLKSDDDTSNPFDEPKNSVWSSPCQLSPSGSGEBLFPVGFYSYKALGILGRSES 433
 DB 1 PFVPTLKSDDDTSNPFDEPKNSVWSSVQQLSPSGSGEBLFPVGFYSYKALGILGRSES 60
 QY 434 VVSGLDSPAKTSMEKKLLIKSKELDSDQKCHKMEQNTRLRRVSEVAIVLSQGEVEL 493
 DB 61 VVSSLDSPAKVSSMEKKLLIKSKELDSDQKCHKMEQNTRLRRVSEVAIVLSQGEVEL 120
 QY 494 KASRTORSLLEODLATYITECSLKRSLQARMEVSOEDDKALQJLHDI REOSRKLOEIK 553
 DB 121 KASRTORSLLEODLATYITECSLKRSLQARMEVSOEDDKALQJLHDI REOSRKLOEIK 180

QY 554 EOEYQAVVEERLMMNOLEEDLVGARRSDLYESELRESRLAAEFKRAKTECOHKLKA 613
 DB 181 EOEYQAVVEERLMMNOLEEDLVGARRSDLYESELRESRLAAEFKRAKTECOHKLKA 240
 QY 614 KDOCKPEVEGEYAKLEKINAEOQLKIOEQLLEKAVKASTEATELLONIRAKERAREL 673
 DB 241 KDOCKPEVEGEYAKLEKINAEOQLKIOEQLLEKAVKASTEATELLONIRAKERAREL 300
 QY 674 EKLONREDSSEKIRKCLVEAEERHSHLENKVRLTEWERRERNRKLDIQTSQOIQWAD 733
 DB 301 EKLHNRDSSSEKIRKCLVEAE----- 321
 QY 734 KILEBEKREAOVSACHLEVHLKQKEQHYEKKIVLDNOIKDLADKETLENNMQRHE 793
 DB 322 ---ELEBEKREAOVSACHLEVHLKQKEQHYEKKIVLDNOIKDLADKETLENNMQRHE 378
 QY 794 EAHKKGKILSBOKMINAMDSKISLEORIVELSEANKLANSSILFORNMKAQEBMISE 853
 DB 379 EAHKKGKILSBOKMINAMDSKISLEORIVELSEANKLANSSILFORNMKAQEBMISE 438
 QY 854 LRQCKFYLETQAGKLEAONRKLIEBQLEKISHQDSKRLLEETRLREVELEHEBQKLE 913
 DB 439 LRQCKFYLETQAGKLEAONRKLIEBQLEKISHQDSKRLLEETRLREVELEHEBQKLE 498
 QY 914 LKRQLTQLSLQERESQTLAQAPALLESQRLQAKTELETTAAEETIOALTARHDE 973
 DB 499 LKRQLTQLSLQERESQTLAQAPALLESQRLQAKTELETTAAEETIOALTARHDE 558
 QY 974 IQRFPDALRNSCTYTTLDEEOLNLTEDNALNNOFTLSQDLBAGANDEIYQLSSEV 1033
 DB 559 IQRFPDALRNSCTYTTLDEEOLNLTEDNALNNOFTLSQDLBAGANDEIYQLSSEV 618
 QY 1034 DILREITEREMOLTQSKOTMEALKTTCTMLEEOVMDLEALNDELLEKEROEAMRSVYG 1093
 DB 619 DILREITEREMOLTQSKOTMEALKTTCTMLEEOVMDLEALNDELLEKEROEAMRSVYG 678
 QY 1094 DEKSQFECRAVELOQMLDTEKQSPARADQRTTESROVVELAVKHEKAIILAOALKEOK 1153
 DB 679 DEKSQFECRAVELOQMLDTEKQSPARADQRTTESROVVELAVKHEKAIILAOALKEOK 738
 QY 1154 LKESLSLKDNDLEKRAMLENNARSIOQKLETERELKORLEEOAKLOOQMDQKXNIF 1213
 DB 739 LKESLSLKDNDLEKRAMLENNARSIOQKLETERELKORLEEOAKLOOQMDQKXNIF 798
 QY 1214 RLTOGLQALDRADLKTESRDELEYOLENIOVLSHEKVMKEGTISOOTKILIDFLQAKMD 1273
 DB 799 RLTOGLQALDRADLKTESRDELEYOLENIOVLSHEKVMKEGTISOOTKILIDFLQAKMD 858
 QY 1274 QPAKKKKVPLQYNELKLALEKARCALEBEALQKTRIELSAREEAHRAKATDHPHST 1333
 DB 859 QPAKKKKVPLQYNELKLALEKARCALEBEALQKTRIELSAREEAHRAKATDHPHST 918
 QY 1334 PATRQOIAMSAIYRSEPHQPSAMSLAPPSSRKESSTPEERRLKERMHNIPRFRN 1393
 DB 919 PATRQOIAMSAIYRSEPHQPSAMSLAPPSSRKESSTPEERRLKERMHNIPRFRN 978
 QY 1394 VGLNMRATKCAVCLDTHFGQASCKLECOVMCHPKSTCIPATCGPAPAYATHTFAFC 1453
 DB 979 VGLNMRATKCAVCLDTHFGQASCKLECOVMCHPKSTCIPATCGPAPAYATHTFAFC 1038
 QY 1454 RDKKNSPGLQTKEPSSILHEGMMKVPNNRKGQOQMDRKITVLEGSKVLIDNEAREAG 1513
 DB 1039 RDKKNSPGLQTKEPSSILHEGMMKVPNNRKGQOQMDRKITVLEGSKVLIDNEAREAG 1098
 QY 1514 QRPVEEBELCLPDGDSVSIHGAVGASELANAKADVPYITLKESHPHRTTCTMPGRTLYILAP 1573
 DB 1099 QRPVEEBELCLPDGDSVSIHGAVGASELANAKADVPYITLKESHPHRTTCTMPGRTLYILAP 1158
 QY 1574 SPFDKQWMTALJESVVGAGVRSREKADALGLNSILKJEGDRLLDMNCTLPPSPDOVYL 1633
 DB 1159 SPFDKQWMTALJESVVGAGVRSREKADALGLNSILKJEGDRLLDMNCTLPPSPDOVYL 1218

QY 1634 VGTEEGYALNVLNLSLTHVPGIAGVFOIYITKLEKLMLAGERACLYDVKKVQSL 1693
 DB 1219 VGTEEGYALNVLNLSLTHVPGIAGVFOIYITKLEKLMLAGERACLYDVKKVQSL 1278
 QY 1694 AOSHLPAPDPISPIFEAVKGCHEGAGKENGCTICAMPKSVILATRENELSKYCTRK 1753
 DB 1279 AOSHLPAPDPISPIFEAVKGCHEGAGKENGCTICAMPKSVILATRENELSKYCTRK 1338
 QY 1754 EITSEPCSIHFTNYSILGTNKFEYIDMKQYTLLEFLDKNDHSLAPAVEAASSNSFPV 1813
 DB 1339 EITSEPCSIHFTNYSILGTNKFEYIDMKQYTLLEFLDKNDHSLAPAVEAASSNSFPV 1398
 QY 1814 SIYVNSAGOREEYLLCFHEGCVVDSYGRSRFTDLKMSLPLAFAYREPLYTTHNS 1873
 DB 1399 SIYVNSAGOREEYLLCFHEGCVVDSYGRSRFTDLKMSLPLAFAYREPLYTTHNS 1458
 QY 1874 LEVIEIQARSAGAPARAYLDIPPRYLGPASSGAYILASSYODKRVLCCKNLVES 1933
 DB 1459 LEVIEIQARSAGAPARAYLDIPPRYLGPASSGAYILASSYODKRVLCCKNLVES 1518
 QY 1934 GTEHHRGPTSRSSPNRKGPTTNEHTTKRVASSPAPPEGSHREBSPHRY--REGRT 1991
 DB 1519 GTEHHRGPTSRSSPNRKGPTTNEHTTKRVASSPAPPEGSHREBSPHRY--REGRT 1578
 QY 1992 ELRRDKSPGRPLEREKSPGRILSTRERSPARLFEDSSRGRLPAGAVRTPLSQNVKWDQ 2051
 DB 1579 ELRRDKSPGRPLEREKSPGRILSTRERSPARLFEDSSRGRLPAGAVRTPLSQNVKWDQ 1638
 QY 2052 SSV 2054
 DB 1639 SSV 1641

RESULT 4

CTRO_MOUSE
 ID CTRO_MOUSE STANDARD; PRT: 1597 AA.
 AC P49025;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 25-OCT-2004 (Rel. 45, Last annotation update)
 DE Citron protein (Rho-interacting, serine/threonine kinase 21).
 GN Name:Cit;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 OC NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=96128238; PubMed=8543060; DOI=10.1016/0014-5793(95)01351-2;
 RA Madanle P., Furuyashiki T., Reid T., Ishizaki T., Watanabe G.,
 RA Morita N., Natsuura S.;
 RT "A novel partner for the GTP-bound forms of rho and rac.";
 RL FEBS Lett. 377:243-248(1995).
 CC -1- FUNCTION: Putative RHO/RAC effector that binds to the GTP-bound
 CC forms of RHO and RAC1. It probably binds p21 with a tighter
 CC specificity in vivo.
 CC -1- SUBUNIT: Homodimer (Probable).
 CC -1- TISSUE SPECIFICITY: A major signal was observed in testis and
 CC brain, but it was also detected in thymus, spleen, kidney, heart
 CC and lung.
 CC -1- SIMILARITY: Contains 1 CNH domain.
 CC -1- SIMILARITY: Contains 1 PH domain.
 CC -1- SIMILARITY: Contains 1 zinc-dependent phospho-ester and DAG
 CC binding domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC -----
DR EMBL; U93904; AAC5234.1; -
DR PIR; S68420; S68420.
DR MGI; MGI:105313; Cit.
DR InterPro; IPR001180; C1tron.
DR InterPro; IPR002219; DAG_PE-bind.
DR InterPro; IPR001849; PH.
DR InterPro; IPR011036; PH_related.
DR InterPro; IPR000861; REM_repeat.
DR Pfam; PF00780; CNH; 1.
DR Pfam; PF00130; DAG_PE-bind; 1.
DR Pfam; PF00169; PH; 1.
DR SMART; SM00109; C1; 1.
DR SMART; SM00036; CNH; 1.
DR SMART; SM00233; PH; 1.
DR PROSITE; PS00479; DAG_PE_BIND_DOM_1; 1.
DR PROSITE; PS00081; DAG_PE_BIND_DOM_2; 1.
DR PROSITE; PS50003; PH_DOMAIN; 1.
KW Coiled coil; Phorbol-ester binding; SH3-binding.
FT DOMAIN 1 845 Coiled coil (Potential).
FT DOMAIN 674 870 Interaction with Rho/Rac.
FT DOMAIN 818 821 Poly-Lys.
FT DOMAIN 931 979 Phorbol-ester and DAG binding.
FT DOMAIN 1011 1131 PH.
FT DOMAIN 1160 1457 CNH.
FT SITE 1521 1526 SH3-binding (Potential).
SQ SEQUENCE 1597 AA; 183448 MW; 787286C2305676DA CRC64;

Query Match 75.6%; Score 7935; DB 1; Length 1597;
Best Local Similarity 98.2%; Pred. No. 8.7e-220;
Matches 1561; Conservative 16; Mismatches 10; Indels 2; Gaps 1;

QY 468 MEOEMTRHRRVSEVAVLSQKEVELKASSETORSILEQDLATYITECSLKRSLQARME 527
DB 9 MEOEMTRHRRVSEVAVLSQKEVELKASSETORSILEQDLATYITECSLKRSLQARME 68
QY 528 VSGEDDKALQLLHDIREQSRKLOEIKOEYQAOYVEEKLMMNOLEEDLVASARRSDLYES 587
DB 69 VSGEDDKALQLLHDIREQSRKLOEIKOEYQAOYVEEKLMMNOLEEDLVASARRSDLYES 128
QY 588 ELRESRLAAEFPRKAEFCQHKLLKADQKPEVGEVAKLEKINABOQAKIQEQLK 647
DB 129 ELRESRLAAEFPRKAEFCQHKLLKADQKPEVGEVAKLEKINABOQAKIQEQLK 188
QY 648 AVKASTATELLONIRAKERAERELEKLNREDSSEGRKCLVEABERHSLKVKYRL 707
DB 189 AVKASTATELLONIRAKERAERELEKLNREDSSEGRKCLVEABERHSLKVKYRL 248
QY 708 ETMERRENRLLKDDIQTKSQOIQWADKILEEKGREAOVSAOHLBVHLKQEQHYEKT 767
DB 249 ETMERRENRLLKDDIQTKSQOIQWADKILEEKGREAOVSAOHLBVHLKQEQHYEKT 308
QY 768 KYLDNQKKKLADKETLENMQRHEEFAHEKGLLSQKAMINAMDSKISLEIRIYELS 827
DB 309 KYLDNQKKKLADKETLENMQRHEEFAHEKGLLSQKAMINAMDSKISLEIRIYELS 368
QY 828 EANKLANSSILFTQRNKKAOEEMISELRQOKFYLETQAGKLEAQNRLKEOLEKISQDH 887
DB 369 EANKLANSSILFTQRNKKAOEEMISELRQOKFYLETQAGKLEAQNRLKEOLEKISQDH 428
QY 888 SDKNRLLELTRLEVSLEHEBQKLEKROLTELQSLQERESQTLAQARALLESQLR 947
DB 429 SDKNRLLELTRLEVSLEHEBQKLEKROLTELQSLQERESQTLAQARALLESQLR 488
QY 948 QAKTELEETTAABEEICALTARDEIQKFDALRNSCTVITDLEBQNLTEENALIN 1007
DB 489 QAKTELEETTAABEEICALTARDEIQKFDALRNSCTVITDLEBQNLTEENALIN 548
QY 1008 QNFYLSKQLEBASGANDEIVQLRSEVDHLRREITEREMQLTISQKQTEALKTCTMLEEQ 1067
DB 549 QNFYLSKQLEBASGANDEIVQLRSEVDHLRREITEREMQLTISQKQTEALKTCTMLEEQ 608
QY 1068 VMDLEALNDELLEKERQWEAMRSVLDGKSGQFECRVRELQRMLDTEKOSRAPADQRTTES 1127

DB 609 VMDLEALNDELLEKERQWEAMRSVLDGKSGQFECRVRELQRMLDTEKOSRAPADQRTTES 668
QY 1128 ROYVELAVKEHKAIEIATLOALKEQKLEKESLSDKNDLEKKAEMENARSIQOKLETE 1187
DB 669 ROYVELAVKEHKAIEIATLOALKEQKLEKESLSDKNDLEKKAEMENARSIQOKLETE 728
QY 1188 RELKORLLEBQALQQOMDLQKXNIFRLTQGLQALDRADLITERSDLEYOLENTQVLY 1247
DB 729 RELKORLLEBQALQQOMDLQKXNIFRLTQGLQALDRADLITERSDLEYOLENTQVLY 788
QY 1248 SHEKVMKEGTISQOTKLIDFLQAKMDQPAKKKVPLOYNELKALKEKARCALEBALQ 1307
DB 789 SHEKVMKEGTISQOTKLIDFLQAKMDQPAKKKVPLOYNELKALKEKARCALEBALQ 848
QY 1308 KTRIELRSAREEAAHKAATDHPSTPATARQOIAMSAIVRSEPHQPSAMSLAPSSRR 1367
DB 849 KTRIELRSAREEAAHKAATDHPSTPATARQOIAMSAIVRSEPHQPSAMSLAPSSRR 908
QY 1368 KESSTPEEFSRRLKEHMHNI PHRFVGLMRAATKCAVCIDTVHFGROASKLECOVMCH 1427
DB 909 KESSTPEEFSRRLKEHMHNI PHRFVGLMRAATKCAVCIDTVHFGROASKLECOVMCH 968
QY 1428 PKCSTCLPATCGLPARYATHTFAFCRDKNSPGLOTKEPSSLHLEGMMKVPRNNRGQ 1487
DB 969 PKCSTCLPATCGLPARYATHTFAFCRDKNSPGLOTKEPSSLHLEGMMKVPRNNRGQ 1028
QY 1488 QGMDRKXIVLEGSKVLIDYNEAREAGORPYEEFELCLPGDVSIGHAVGASELANTAKAD 1547
DB 1029 QGMDRKXIVLEGSKVLIDYNEAREAGORPYEEFELCLPGDVSIGHAVGASELANTAKAD 1088
QY 1548 VPYILKMEHPHTTQCPGRTLYLLAPSFPDKQKQWTVALESVAVGVRSEKAEADAKLIG 1607
DB 1089 VPYILKMEHPHTTQCPGRTLYLLAPSFPDKQKQWTVALESVAVGVRSEKAEADAKLIG 1148
QY 1608 NSLLKLEGGDRLLMNCITLPPSDQVLTGTEEGYALNVLKNSLTHIPGIAVFOIYIKD 1667
DB 1149 NSLLKLEGGDRLLMNCITLPPSDQVLTGTEEGYALNVLKNSLTHIPGIAVFOIYIKD 1208
QY 1668 LEKILMAGEERALCLVDYVKVQSLAQSHLPAPQDISPNIPEAVKCHLFGAGKLENGL 1727
DB 1209 LEKILMAGEERALCLVDYVKVQSLAQSHLPAPQDISPNIPEAVKCHLFGAGKLENGL 1268
QY 1728 CICAAMPKVVILRYNEMISKCIKRIETSEPCSCIHFTNYSILIGTNKFEYIDMKQYT 1787
DB 1269 CICAAMPKVVILRYNEMISKCIKRIETSEPCSCIHFTNYSILIGTNKFEYIDMKQYT 1328
QY 1788 LEEFLDKNDHSLAPVPAASNSFPVSIYQVNSAGQREEVLLCFHEGFVVDYGRSRRT 1847
DB 1329 LEEFLDKNDHSLAPVPAASNSFPVSIYQVNSAGQREEVLLCFHEGFVVDYGRSRRT 1388
QY 1848 DDLMKSRLLPLAFYREPYLFTVHFNSLEVEIQARSSAGTPARAYLIDIPNRYLGPALIS 1907
DB 1389 DDLMKSRLLPLAFYREPYLFTVHFNSLEVEIQARSSAGTPARAYLIDIPNRYLGPALIS 1448
QY 1908 GAIYLLASVQDKRLVICCKGNLVKESGTEHHRGPSTSRSSPNRRGPPTYNEHTTKVASS 1967
DB 1449 GAIYLLASVQDKRLVICCKGNLVKESGTEHHRGPSTSRSSPNRRGPPTYNEHTTKVASS 1508
QY 1968 PAPPEGSHRREESTHRY--REGRTLRDQKSPGPLEEKSPPGILLRREERSPARLF 2025
DB 1509 PAPPEGSHRREESTHRY--REGRTLRDQKSPGPLEEKSPPGILLRREERSPARLF 1568
QY 2026 EDSRGRPLPAGAVRTPLSQVNAKWDOSV 2054
DB 1569 EDSRGRPLPAGAVRTPLSQVNAKWDOSV 1597

RESULT 5
ID 090X19 PRELIMINARY; PRT; 1618 AA.
AC 090X19;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)

DT 01-MAY-2000 (Tremblrel. 13, last sequence update)
 DT 01-MAY-2004 (Tremblrel. 26, last annotation update)
 DE Postsynaptic density protein.
 GN Name=Citron;
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
 NX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Zhang W., Apperson M.L., Vasquez L.E., Kennedy M.B.;
 RL Submitted (DEC-1997) to the EMBL/Genbank/DBJ databases.
 CC -1- SIMILARITY: Contains 1 PH domain.
 DR EMBL; AF039218; AAC25483.1; -.
 DR GO; GO:0005622; C:intracellular; IEA.
 DR GO; GO:0005083; F:small GTPase regulatory/interacting protein. .; IEA.
 DR GO; GO:0007242; P:intracellular signaling cascade; IEA.
 DR InterPro; IPR001180; Citron.
 DR InterPro; IPR005479; Gphp_synth_L_D2.
 DR InterPro; IPR002219; DAG_pe_bind.
 DR InterPro; IPR001849; PH.
 DR InterPro; IPR000861; REM_repeat.
 DR InterPro; IPR000533; Tropomyosin.
 DR Pfam; PF00130; Cl_1; 1.
 DR Pfam; PF00780; CNH; 1.
 DR Pfam; PF00169; PH; 1.
 DR PRINTS; PR00194; TROPOMYOSIN.
 DR SMART; SMO0109; Cl_1.
 DR SMART; SMO0036; CNH; 1.
 DR SMART; SMO0233; PH; 1.
 DR PROSITE; PS00867; CPSASE_2; UNKNOWN_1.
 DR PROSITE; PS00479; DAG_PE_BIND_DOM_1; 1.
 DR PROSITE; PS50081; DAG_PE_BIND_DOM_2; 1.
 DR PROSITE; PS50003; PH_DOMAIN; 1.
 SQ SEQUENCE 1618 AA; 185924 MW; EAEDE079788AA4E9 CRC64;

Query Match 74.8%; Score 7851.5; DB 2; Length 1618;
 Best Local Similarity 96.0%; Pred. No. 2.2e-217;
 Matches 1554; Conservative 18; Mismatches 14; Indels 33; Gaps 4;

QY 468 MEQEMTLHRRVSEVAVLSQKEVELKASEFORSLLSODLATYTTESSLKRSLEQARME 527
 DB 1 MEQEMTLHRRVSEVAVLSQKEVELKASEFORSLLSODLATYTTESSLKRSLEQARME 60
 QY 528 VSGEDDVALQLHDIRESQRKLQEIKEQVQAYEENKLMNQLEEDLVARSRSDIYES 587
 DB 61 VSGEDDVALQLHDIRESQRKLQEIKEQVQAYEENKLMNQLEEDLVARSRSDIYES 120
 QY 588 ELRESRLAAEEFKKATECOHKLKAKDOCKPVEGEYAKLEKINAEOQLKIQELQEKLEK 647
 DB 121 ELRESRLAAEEFKKATECOHKLKAKDOCKPVEGEYAKLEKINAEOQLKIQELQEKLEK 180
 QY 648 AVKASTATLQNIROAKERAERLEKLNREDSSGIRKCLVYEAERRLSENKTKYRL 707
 DB 181 AVKASTATLQNIROAKERAERLEKLNREDSSGIRKCLVYEAERRLSENKTKYRL 240
 QY 708 ETMERRENRLKDDIQTSQOIQOMADKILLEEKGREAOVSAOHL EYHLKQKEQYHEKI 767
 DB 241 ETMERRENRLKDDIQTSQOIQOMADKILLEEKGREAOVSAOHL EYHLKQKEQYHEKI 300
 QY 768 KVLNDNQIKDLADKETLNNMQRHEEENAEKGLISQOKAMINAMDSKIRSLRQRIYELS 827
 DB 301 KVLNDNQIKDLADKETLNNMQRHEEENAEKGLISQOKAMINAMDSKIRSLRQRIYELS 360
 QY 828 EANTLANSSLPTRNNKAQOEMISELRQCKFYLETQAGKLEAONRKLEBOLKISQDH 887
 DB 361 EANTLANSSLPTRNNKAQOEMISELRQCKFYLETQAGKLEAONRKLEBOLKISQDH 420
 QY 888 SSKRLLELETRLEVSLEHEQKLELKRQUTLEQLSIOBRESQUTLQAARALLESOLR 947
 DB 421 SSKRLLELETRLEVSLEHEQKLELKRQUTLEQLSIOBRESQUTLQAARALLESOLR 480
 QY 948 QAKTELEETTAABEEIQAALTAHDEIQRKFDALRNSCTVITDLEQLNQTEDNAELNN 1007

DB 481 QAKTELEETTAABEEIQAALTAHDEIQRKFDALRNSCTVITDLEQLNQTEDNAELNN 540
 QY 1008 QNFYLSKQLDASGANDEIVOLRSEVDHLREITEREMQWLSQKQTEALKTCTMLEEQ 1067
 DB 541 QNFYLSKQLDASGANDEIVOLRSEVDHLREITEREMQWLSQKQTEALKTCTMLEEQ 600
 QY 1068 VMDLEALNDELLEKEREKOWEAWRSVLDGKSGQFECRVRELQRMIDTERKQSRARADQRTES 1127
 DB 601 VMDLEALNDELLEKEREKOWEAWRSVLDGKSGQFECRVRELQRMIDTERKQSRARADQRTES 660
 QY 1128 ROVVELAVKHKEEITLQALKEQKIKASLSQKNDLEKHAMLENNAISLQOKETE 1187
 DB 661 ROVVELAVKHKEEITLQALKEQKIKASLSQKNDLEKHAMLENNAISLQOKETE 720
 QY 1188 RELKORLLEBOAKLQOQMDLQKNI FRLTQGLQBALRADLKTERSDLEQLENIQVLY 1247
 DB 721 RELKORLLEBOAKLQOQMDLQKNI FRLTQGLQBALRADLKTERSDLEQLENIQVLY 780
 QY 1248 SHEKVKMEGTISQOTKLIDFLQAMQDPAKKKK-----VPLQVNEKLAL 1292
 DB 781 SHEKVKMEGTISQOTKLIDFLQAMQDPAKKKKLFRRKEDPALPQVPLQVNEKLAL 840
 QY 1293 EKEKARCAELEALQKTRIELRSAREAAHRAKATDHPSTPATARQOIMSAIVRSPEH 1352
 DB 841 EKEKARCAELEALQKTRIELRSAREAAHRAKATDHPSTPATARQOIMSAIVRSPEH 900
 QY 1353 QPSAMSLAPPSRRKESSTPEEFSRLKERMHNI PHRFVNGAMBATKACVCLDTVHF 1412
 DB 901 QPSAMSLAPPSRRKESSTPEEFSRLKERMHNI PHRFVNGAMBATKACVCLDTVHF 960
 QY 1413 GROSXKLECOVMCHPCSTCLPATCGLPAEVATHTFAECRDQNSPGLOTKEPSSSLH 1472
 DB 961 GROSXKLECOVMCHPCSTCLPATCGLPAEVATHTFAECRDQNSPGLOTKEPSSSLH 1020
 QY 1473 LEGMMKTPRNKKGQOQMDKRYLYLESGSKYLIYDNEAREAGQRPVEFELCPFGDVSIH 1532
 DB 1021 LEGMMKTPRNKKGQOQMDKRYLYLESGSKYLIYDNEAREAGQRPVEFELCPFGDVSIH 1080
 QY 1533 GAVGASELATYAKADVYIILKMEHPHTTCMPGRTYLLAPSPFDKRWVTALSSVAVAG 1592
 DB 1081 GAVGASELATYAKADVYIILKMEHPHTTCMPGRTYLLAPSPFDKRWVTALSSVAVAG 1140
 QY 1593 RVSRKAEADA-----KLGNSLLKLEGDDRLDMNCTLPSGDDVVLVGE 1637
 DB 1141 RVSRKAEADAANDCTSCEBLPVWEKLLGNSLLKLEGDDRLDMNCTLPSGDDVVLVGE 1200
 QY 1638 EGLYALNVLKNSLTHVPGIGAVFOIYIILKLEKLMJAGEBRALCLVDYKVKOSLAOSH 1697
 DB 1201 EGLYALNVLKNSLTHVPGIGAVFOIYIILKLEKLMJAGEBRALCLVDYKVKOSLAOSH 1260
 QY 1698 LPAPODISPIPEAVKGCHEFGAGKLENGLCIQAAMSKVYILRYNNLSKYCIKKKEET 1757
 DB 1261 LPAPODISPIPEAVKGCHEFGAGKLENGLCIQAAMSKVYILRYNNLSKYCIKKKEET 1320
 QY 1758 SEPSCIHFNYSILIGTNKFYEIDMKQYTLBEFLDKNDLSLAPAVEAASNSPFSVIVQ 1817
 DB 1321 SEPSCIHFNYSILIGTNKFYEIDMKQYTLBEFLDKNDLSLAPAVEAASNSPFSVIVQ 1380
 QY 1818 VNSAGQREBYLLCFHEGFVDSYGRSRITDMLKMSRLPLAFAYREBYLLPYTHNSLEVI 1877
 DB 1381 ANSTGQREBYLLCFHEGFVDSYGRSRITDMLKMSRLPLAFAYREBYLLPYTHNSLEVI 1440
 QY 1878 EIQARSSAGTPARAYLIDIPNRYGPAISSGAYILASSYODKLRVITCKNGLVESGTEH 1937
 DB 1441 EIQARSSAGTPARAYLIDIPNRYGPAISSGAYILASSYODKLRVITCKNGLVESGTEH 1500
 QY 1938 HRGPSTSRSPNKGAPTYNEHTTKRVASSPAPPEGSHREBSPHRY--REGTELR 1995
 DB 1501 HRGPSTSR--SPNKGAPTYNEHTTKRVASSPAPPEGSHREBSPHRYDRBERTELRR 1559
 QY 1996 DKSPGRLEBEKAPGRLSTRRERSPARLFEDESRGLPAGAVTPLSGVNNKVDGSSV 2054

Db 1560 DKSPGRLEHKEKSPGRMLSTRERSPRLFEEDSSRGRLPAGAVRTPLSQVNVKWDQSSV 1618

RESULT 6
CTRO HUMAN
ID CTRO HUMAN STANDARD; PRT, 1266 AA.
AC 014578; 09UP27;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Cytro protein (Rho-interacting, serine/threonine kinase 21) (Fragment).
GN Name=CIT; Synonyms=KIA0949, STR21;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Comell M., Goela D., Harper M.,
RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE OF 347-1286 FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=99246063; PubMed=10231032;
RA Nagaie T., Ishikawa K.-I., Suyama M., Kikuno R., Hirosewa M.,
RA Miyajima N., Tanaka A., Kotani H., Nomura N., Ohara O.,
RT "Prediction of the coding sequences of unidentified human genes. XIII.
RT The complete sequences of 100 new cDNA clones from brain which code
RT for large proteins in vitro.",
RL DNA Res. 6:63-70(1999).
CC -1- FUNCTION: Putative RHO/RAC effector that binds to the GTP-bound
CC forms of RHO and RAC1. It probably binds p21 with a tighter
CC specificity in vivo (By similarity).
CC -1- SIMILARITY: Contains 1 CNH domain.
CC -1- SIMILARITY: Contains 1 PH domain.
CC -1- SIMILARITY: Contains 1 zinc-dependent phorbol-ester and DAG
CC binding domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by, and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AC002563; AAB71327.1; -.
CC EMBL: AB023166; BAA76793.1; -.
CC Genew: HSNCL1985; CIT.
DR MIM: 605629; -.
DR InterPro: IPR001180; Cytro.
DR InterPro: IPR002219; DAG_PE-bind.
DR InterPro: IPR001849; PH_repeat.
DR InterPro: IPR01036; PH_related.
DR InterPro: IPR000861; KEM_repeat.
DR Pfam: PF00780; CNH; 1.
DR Pfam: PF00130; DAG_PE-bind; 1.
DR Pfam: PF00169; PH; 1.
DR SMART: SM00109; C1; 1.
DR SMART: SM00036; CNH; 1.
DR SMART: SM00233; PH; 1.
DR PROSITE: PS00479; DAG_PE_BIND_DOM_1; 1.
DR PROSITE: PS50081; DAG_PE_BIND_DOM_2; 1.
DR PROSITE: PS50003; PH_DOMAIN; 1.
KW Coiled coil; Phorbol-ester binding; SH3-binding.
FT NON_TER 1
FT DOMAIN <1 558 Coiled coil (Potential).
FT DOMAIN 365 561 Rho/Rac binding.
FT DOMAIN 509 512 Poly-Lys.
FT DOMAIN 622 670 Phorbol-ester and DAG binding.
FT DOMAIN 702 822 PH.
FT DOMAIN 851 1148 CNH.

FT SITE 1212 1217 SH3-binding (Potential).
SQ SEQUENCE 1266 AA; 146506 MM; 498101F79EA75E85 CAC64;
Query Match 62.7%; Score 6575; DB 1; Length 1266;
Best Local Similarity 99.8%; Pred. No. 6,3e-181;
Matches 1284; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 769 VLNDNQIKDLADKETLNNMORHEEFNAHEKGLTSBOKAMINMDSKIRSLBORIVLSE 828
DB 1 VLNDNQIKDLADKETLNNMORHEEFNAHEKGLTSBOKAMINMDSKIRSLBORIVLSE 60
QY 829 ANKLANSLSFTGRNKAQEMISELRQKFTYLETQAGKLEAQRKLEQLERKISHQDS 888
DB 61 ANKLANSLSFTGRNKAQEMISELRQKFTYLETQAGKLEAQRKLEQLERKISHQDS 120
QY 889 DKNRLLETRLEVSLEHEEQKLEKRLQTELQLSIQERESQTLQAARALLESQLRQ 948
DB 121 DKNRLLETRLEVSLEHEEQKLEKRLQTELQLSIQERESQTLQAARALLESQLRQ 180
QY 949 AKTELEETTAEEAEIQLALTARDEIQKFDALRNSCTVITDLEEQNLQTEDNAELNQ 1008
DB 181 AKTELEETTAEEAEIQLALTARDEIQKFDALRNSCTVITDLEEQNLQTEDNAELNQ 240
QY 1009 NFYLRSQDLDEASGANDEIVQLRSEVDHLRREITEREMQTSQKQMEALKTCTMLEEQV 1068
DB 241 NFYLRSQDLDEASGANDEIVQLRSEVDHLRREITEREMQTSQKQMEALKTCTMLEEQV 300
QY 1069 MDLEALNDELLEKEROMEANRSLVGDSEKQFECRAVELQMDLTERKOSRARDQRTESR 1128
DB 301 MDLEALNDELLEKEROMEANRSLVGDSEKQFECRAVELQMDLTERKOSRARDQRTESR 360
QY 1129 QVVELAVKEKKAITLQALKEQKIKASLSKNDLEKHAMLEMNARSLOQKLETER 1188
DB 361 QVVELAVKEKKAITLQALKEQKIKASLSKNDLEKHAMLEMNARSLOQKLETER 420
QY 1189 ELKORLLEBAKIQOQMDLQKNIIFRLTQGLABLRADLTKTERSDLEXYOLENIQVLYS 1248
DB 421 ELKORLLEBAKIQOQMDLQKNIIFRLTQGLABLRADLTKTERSDLEXYOLENIQVLYS 480
QY 1249 HEKVKMEGTISQOTKLIDFLQAMDQPAKKKQVPLQYNELKALEKKAQCAELEALQK 1308
DB 481 HEKVKMEGTISQOTKLIDFLQAMDQPAKKKQVPLQYNELKALEKKAQCAELEALQK 540
QY 1309 TRIELASAREAAHRAKATDHPHSTPATATARQOQJAMAAIYSPHOSAMSLAPSSRRK 1368
DB 541 TRIELASAREAAHRAKATDHPHSTPATATARQOQJAMAAIYSPHOSAMSLAPSSRRK 600
QY 1369 ESSTPEFSRRLKERMHNIIPHFENVGLMNRATKCAVCLDTYAFGRQASKLEQVWCHP 1428
DB 601 ESSTPEFSRRLKERMHNIIPHFENVGLMNRATKCAVCLDTYAFGRQASKLEQVWCHP 660
QY 1429 KCSTCLPATCGLPAEYATHTFAFCRDKNMSPGLQTKPESSLAHEGMKVPNNRGGQ 1488
DB 661 KCSTCLPATCGLPAEYATHTFAFCRDKNMSPGLQTKPESSLAHEGMKVPNNRGGQ 720
QY 1489 GMDRKTYIVLEGSVYLINDNARARAGRPVEEFLCPDGDVYHGAVGASLANTAKADV 1548
DB 721 GMDRKTYIVLEGSVYLINDNARARAGRPVEEFLCPDGDVYHGAVGASLANTAKADV 780
QY 1549 PYLKNESHPTTQWPGRTYTLAPSPDPQKQVATLLESVAVAGARVREKAEADAKLGN 1608
DB 781 PYLKNESHPTTQWPGRTYTLAPSPDPQKQVATLLESVAVAGARVREKAEADAKLGN 840
QY 1609 SLILKGBDRDLNNCTPLPSGDQVVLVGTBEGALVALNVLKNSLTHVIGAVFOYIYIKDL 1668
DB 841 SLILKGBDRDLNNCTPLPSGDQVVLVGTBEGALVALNVLKNSLTHVIGAVFOYIYIKDL 900
QY 1669 EKLIMTAGEERACLVDVKKVKOSLAQSHLPAQPDISPNIPEAVKCGHLFGAKITENGUC 1728
DB 901 EKLIMTAGEERACLVDVKKVKOSLAQSHLPAQPDISPNIPEAVKCGHLFGAKITENGUC 960
QY 1729 ICAAMPKVVILLYNNLSKYCIKKEITENSEPSCSHFNTYSTLIGNKPYEIDMKQYTL 1788

Db 961 ICAMPKVVILARNENLSKYCIKRIETSEPCSIHFTNYSILGTNKFYEIDMKOYTL 1020
 QY 1789 EEPFDKNDHSLAPVPAASNSPFSVIQVNSAGOREEYLLCFHEFGVFDVSYGRSRSTD 1848
 Db 1021 EEPFDKNDHSLAPVPAASNSPFSVIQVNSAGOREEYLLCFHEFGVFDVSYGRSRSTD 1080
 QY 1849 DLKKSRLPLAFVAREPFLVTHFNSLEVIETQARSSACTPRAVLDINPRTYLGPAISSG 1908
 Db 1081 DLKKSRLPLAFVAREPFLVTHFNSLEVIETQARSSACTPRAVLDINPRTYLGPAISSG 1140
 QY 1909 AIYLAASVQDRLARYICCKGNLVKESGTEHHRGSTSRSSPNKRGPPTYNEHTKRVASSP 1968
 Db 1141 AIYLAASVQDRLARYICCKGNLVKESGTEHHRGSTSRSSPNKRGPPTYNEHTKRVASSP 1200
 QY 1969 APPEGSPHREPSPPHRYREGRTETLRDSDPGRPLERKSPGRIILSTRERSPARLPEDS 2028
 Db 1201 APPEGSPHREPSPPHRYREGRTETLRDSDPGRPLERKSPGRIILSTRERSPARLPEDS 1260
 QY 2029 SRGRLPAGAVTPTPLSQVKNKWDQSSV 2054
 Db 1261 SRGRLPAGAVTPTPLSQVKNKWDQSSV 1286
 RESULT 7
 ID Q8CIJ1 PREDIMINARY; PRT; 1032 AA.
 AC Q8CIJ1;
 DT 01-MAR-2003 (TREMblrel. 23, Created)
 DT 01-MAR-2003 (TREMblrel. 23, Last sequence update)
 DT 05-JUL-2004 (TREMblrel. 27, Last annotation update)
 DE C1c protein.
 GN Name=C1c;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 CX NCBI_TaxID=10090;
 RN [1]
 RC SEQUENCE FROM N.A.
 RC STRAIN=FVB/N; TISSUE=Mammary tumor. C3;
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strauberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shemmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Ueda T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Rana S.S., Loquellano N.A., Peters G.J., Adamson R.D., Mullany S.J.,
 RA Bosak S.A., McGowan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hultik S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butcherfield Y.S.,
 RA Krzyzanski M.I., Skalska U., Smallus D.E., Scherch A., Schein J.E.,
 RA Jones S.J., Marra M.A.;
 RT "generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [2]
 RC SEQUENCE FROM N.A.
 RC STRAIN=FVB/N; TISSUE=Mammary tumor. C3;
 RA Strauberg R.; TISSUE=Mammary tumor. C3;
 RN [3]
 RC SEQUENCE FROM N.A.
 RC STRAIN=FVB/N; TISSUE=Mammary tumor. C3;
 RA Strauberg R.;
 RL Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: Contains 1 PH domain.
 DR EMBL/ BC023775; AAH23775.1; -.
 DR EMBL/ BC051165; AAH51165.1; -.

DR MED, MG1.105313; C1c.
 DR GO; GO:0005083; F:small GTPase regulatory/interacting protein. . .; IEA.
 DR GO; GO:0007242; P:intracellular signaling cascade; IEA.
 DR InterPro; IPR001180; C1c1on.
 DR InterPro; IPR005479; C1c1n synch L_D2.
 DR InterPro; IPR002219; DAG_PE-bind.
 DR InterPro; IPR001849; PH.
 DR InterPro; IPR011036; PH-related.
 DR Pfam; PF00130; C1_1; 1.
 DR Pfam; PF00780; CNH; 1.
 DR Pfam; PF00169; PH; 1.
 DR SMART; SM00109; C1; 1.
 DR SMART; SM00036; CNH; 1.
 DR SMART; SM00233; PH; 1.
 DR PROSITE; PS00867; CPSASE_2; UNKNOWN 1.
 DR PROSITE; PS00479; DAG_PE_BIND_DOM_1; 1.
 DR PROSITE; PS50081; DAG_PE_BIND_DOM_2; 1.
 DR PROSITE; PS50003; PH_DOMAIN; 1.
 DR SEQUENCE 1032 AA; 116918 MW; 59ABDC9659DA33B CRC64;
 Query Match 48.2%; Score 5057.5; DB 2; Length 1032;
 Best Local Similarity 95.1%; Pred. No. 1.5e-137;
 Matches 982; Conservative 10; Mismatches 8; Indels 33; Gaps 4;
 QY 1054 MEALKTCTMLEBOVMDLEALNDELKERQWEAMRSVLDGKSGQFGRVRELQRMIDTE 1113
 Db 1 MEALKTCTMLEBOVMDLEALNDELKERQWEAMRSVLDGKSGQFGRVRELQRMIDTE 60
 QY 1114 KQSRARADQRTESROVVELAVKHKAEILLALQALKEQKAKASLSDKNDLDEKKAHL 1173
 Db 61 KQSRARADQRTESROVVELAVKHKAEILLALQALKEQKAKASLSDKNDLDEKKAHL 120
 QY 1174 ENMARSLOQKLETRERELKORLLEBOAKLQOQMDLOKNHIFLITQGLQALDRADLKTET 1233
 Db 121 ENMARSLOQKLETRERELKORLLEBOAKLQOQMDLOKNHIFLITQGLQALDRADLKTET 180
 QY 1234 SDLEYOLENITQVLYSHKVKVMGGTISQOTKILIDFLQAKMDQPAKKK----- 1280
 Db 181 SDLEYOLENITQVLYSHKVKVMGGTISQOTKILIDFLQAKMDQPAKKK----- 1280
 QY 1281 --VLOVNEKLALKEKKARCALEELQKTRIELRSAREBAARKATDHPSTPATAR 1338
 Db 241 TVPLQVNEKLALKEKKARCALEELQKTRIELRSAREBAARKATDHPSTPATAR 300
 QY 1339 QQIMSAIVSPHEQIPAMSLAPSSRRKSSSTPEESRRLKERMHNHIFRNVGLAM 1398
 Db 301 QQIMSAIVSPHEQIPAMSLAPSSRRKSSSTPEESRRLKERMHNHIFRNVGLAM 360
 QY 1399 RATKCAVCLDTVHFGROASKLECOVCHPCKSTCLPATGCLPAEVATHTFEACRDRGN 1458
 Db 361 RATKCAVCLDTVHFGROASKLECOVCHPCKSTCLPATGCLPAEVATHTFEACRDRGN 420
 QY 1459 SPGLQTEPSSSLHLEGKMKYPRNNKGGQGMKRYIVLBSKYLITDNEAREKQRPVE 1518
 Db 421 SPGLQTEPSSSLHLEGKMKYPRNNKGGQGMKRYIVLBSKYLITDNEAREKQRPVE 480
 QY 1519 EPELCLPDGDSYIHGAAGASLANTARADVYTIKMSHPHTTCGPGRTYLLAPSPDX 1578
 Db 481 EPELCLPDGDSYIHGAAGASLANTARADVYTIKMSHPHTTCGPGRTYLLAPSPDX 540
 QY 1579 QRWTALESVVGAGVRERKAADA-----KLGNLSLKLKGDRDLDMNC 1623
 Db 541 QRWTALESVVGAGVRERKAADAADHTSSEHQPVWEKTLGNLSLKLKGDRDLDMNC 600
 QY 1624 TLPPSDQVTVGTBEGLYALNVLKNSLTHVPGICAVQIYIYIKOLEKILMAGEBALCL 1683
 Db 601 TLPPSDQVTVGTBEGLYALNVLKNSLTHVPGICAVQIYIYIKOLEKILMAGEBALCL 660
 QY 1684 VDVKKVQSLAQSHLPAOPDISPVIPEAVKGCFLFGAGKTEGLGCTICAMPKSVVILARN 1743
 Db 661 VDVKKVQSLAQSHLPAOPDISPVIPEAVKGCFLFGAGKTEGLGCTICAMPKSVVILARN 720
 QY 1744 ENLSKYCIKRIETSEPCSIHFTNYSILGTNKFYEIDMKOYTLBEFLDKNDHSLAPAV 1803


```

Db 721 DNASKYCIKREIETSEPOSCIHFTNYSILGTNFTYIDMKQYTLDEFLLDKNDHSLAPV 780
QY 1804 FAASNSFPVSIYVONAGOREEYLGFHEGFVDSYGRSRITDIDKMSRLPLAFAYRE 1863
Db 781 FASSNSFPVSIYVONAGOREEYLGFHEGFVDSYGRSRITDIDKMSRLPLAFAYRE 840
QY 1864 PYLFTVTHNSLVEIETIQARSSAGCPAPAYLDIPNRYLGAIPSSGAIYLAASYODKLRYI 1923
Db 841 PYLFTVTHNSLVEIETIQARSSAGCPAPAYLDIPNRYLGAIPSSGAIYLAASYODKLRYI 900
QY 1924 CCKGNLVKESGTEHHRPSTSRSSPNKRGPTVNEHTTKRYASSPAPREGSHREPSTP 1983
Db 901 CCKGNLVKESGTEHHRPSTSRSSPNKRGPTVNEHTTKRYASSPAPREGSHREPSTP 959
QY 1984 HRY--REGRTELRRDKSPGRPLEREKSPGRILSTRERSPARLFEDESRCGLPAGAVRTP 2041
Db 960 HRYADRGRTELRRDKSPGRPLEREKSPGRMLSTRERSPARLFEDESRCGLPAGAVRTP 1019
QY 2042 LSGVNVKWDQSSV 2054
Db 1020 LSGVNVKWDQSSV 1032

```

RESULT 8

```

Q6XUH8 PRELIMINARY; PRT; 482 AA.
AC Q6XUH8;
DT 05-JUL-2004 (TReMBLrel. 27, Created)
DT 05-JUL-2004 (TReMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TReMBLrel. 27, Last annotation update)
DE Citron Rho-interacting kinase short form.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Mao Y., Xie Y., Wu Q.;
RC Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Belongs to the Ser/Thr protein kinase family.
DE EMBL: AY209000; AAF43922.1; -.
DR HSSP; P05132; IATP.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR GO; GO:0004713; F:protein-tyrosine kinase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR InterPro; IPR011009; Kinase_lk.
DR InterPro; IPR000961; Kinase_C.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR002290; Ser_Thr_kinase.
DR InterPro; IPR008271; Ser_Thr_kinase.
DR InterPro; IPR01245; Tyr_kinase.
DR Pfam; PF00069; Kinase_C; 1.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00220; S_TKc_1.
DR SMART; SM00133; S_TK_X; 1.
DR SMART; SM00219; TyKc; 1.
DR PROSITE; PS00107; PROTEIN KINASE_ATP; 1.
DR PROSITE; PS50011; PROTEIN KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN KINASE_ST; 1.
KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 482 AA; 5453 MW; 4F02AF6A32C73BAE CRC64;

```

```

Query Match 23.7%; Score 2492; DB 2; Length 482;
Best Local Similarity 99.4%; Pred. No. 2-2e-64;
Matches 477; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

```

```

QY 1 MLKRYGARNPLDGAAPLPIASRSRLNLFQKGPPTMTQOMSPLSREGILDALFVLF 60
Db 1 MLKRYGARNPSDGAAPLPIASRSRLNLFQKGPPTMTQOMSPLSREGILDALFVLF 60

```

```

QY 61 ECGPAMLKIKVSNTRKXSDTIAELQELQPSANDFEVRSLYGCGHFAVQVVRKATG 120
Db 61 ECGPAMLKIKVSNTRKXSDTIAELQELQPSAKOFEVRSLYGCGHFAVQVVRKATG 120
QY 121 DIYAMVKKKKALLAEOVSFFFEERNILSRSPMIPOLQYAFODKNHLYLWMEYOPGG 180
Db 121 DIYAMVKKKKALLAEOVSFFFEERNILSRSPMIPOLQYAFODKNHLYLWMEYOPGG 180
QY 181 DLASLNRVEDQDENLQFYLAELILAVSHVMGVYHVDIKPENILVDRTGHIKLVDF 240
Db 181 DLASLNRVEDQDENLQFYLAELILAVSHVMGVYHVDIKPENILVDRTGHIKLVDF 240
QY 241 GSAAKNSNMVNAKPIGTPDYMADEVLTVWNGDGKGTGLDCDWMSVGVIAVEMTYGR 300
Db 241 GSAAKNSNMVNAKPIGTPDYMADEVLTVWNGDGKGTGLDCDWMSVGVIAVEMTYGR 300
QY 301 SPPAEGTSARTENNINNFQELFKPPDPKVSDFLLIOSLIGOKERLKEGICCHPFF 360
Db 301 SPPAEGTSARTENNINNFQELFKPPDPKVSDFLLIOSLIGOKERLKEGICCHPFF 360
QY 361 SKIDMNNIRNSPPFPYTLKSDDTNSFDEPEKNSWSSPCQLSPSGFSGEELPFVGF 420
Db 361 SKIDMNNIRNSPPFPYTLKSDDTNSFDEPEKNSWSSPCQLSPSGFSGEELPFVGF 420
QY 421 YSRALGILGRSESVSGLDSPATSSMEKKLILKSYELQDSQDKCHMBOEMTRLIRVS 480
Db 421 YSRALGILGRSESVSGLDSPATSSMEKKLILKSYELQDSQDKCHMBOEMTRLIRVS 480

```

RESULT 9

```

O88937 PRELIMINARY; PRT; 494 AA.
AC O88937;
DT 01-NOV-1998 (TReMBLrel. 08, Created)
DT 01-NOV-1998 (TReMBLrel. 08, Last sequence update)
DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)
DE Rho/rac-interacting citron kinase short isoform.
GN Name=Cit; Synonyms=Crik-sk;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99009084; Pubmed=9792683; DOI=10.1074/jbc.273.45.29706;
RA Di Cunto F., Calautti E., Heisio J., Ong L., Topley G., Turco E.,
RA Decto G.P.;
RT "Citron rho-interacting kinase, a novel tissue-specific ser/thr kinase
RT encompassing the Rho-Rac-binding protein Citron."
RL J. Biol. Chem. 273:29706-29711(1998).
CC -1- SIMILARITY: Belongs to the Ser/Thr protein kinase family.
DE EMBL: AF086823; AAC72822.1; -.
DR HSSP; P31751; IMRV.
DR MCD; MG1105313; Cit.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR GO; GO:0004740; F:transferase activity; IEA.
DR GO; GO:0016740; P:protein amino acid phosphorylation; IEA.
DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR InterPro; IPR011009; Kinase_lk.
DR InterPro; IPR000961; Kinase_C.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR002290; Ser_Thr_kinase.
DR InterPro; IPR008271; Ser_Thr_kinase.
DR Pfam; PF00069; Kinase_C; 1.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00220; S_TKc_1.
DR SMART; SM00133; S_TK_X; 1.
DR PROSITE; PS00107; PROTEIN KINASE_ATP; 1.
DR PROSITE; PS50011; PROTEIN KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN KINASE_ST; 1.
KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.

```

SQ SEQUENCE 494 AA; 55742 MW; 2C60D2294B28A1B5 CRC64;
 Query Match 20.7%; Score 2175.5; DB 2; Length 494;
 Best Local Similarity 88.2%; Pred. No. 2.6e-55;
 Matches 413; Conservative 25; Mismatches 29; Indels 1; Gaps 1;
 QY 1 MLKRYGARNPLDGAAPIASRASRLNLFQGGKPPMTQOQMSPLSREGILDALFVLE 60
 DB 1 MLKRYGARNPLDGAAPIASRASRLNLFQGGKPPMTQOQMSPLSREGILDALFVLE 60
 QY 61 ECSOPALMKIKHVSNFVKYSDTIAEIOEPSAKDFEVSILVCGHFAEYQVVERKATG 120
 DB 61 ECSOPALMKIKHVSNFVKYSDTIAEIOEPSAKDFEVSILVCGHFAEYQVVERKATG 120
 QY 121 DIYAMKVMKKKALLAEOQVSPFEEERNILSRSTSPMIPOLOAFQDKNHLVMEYQPG 180
 DB 121 DIYAMKVMKKKALLAEOQVSPFEEERNILSRSTSPMIPOLOAFQDKNHLVMEYQPG 180
 QY 121 DVYAMKIMKKKALLAEOQVSPFEEERNILSRSTSPMIPOLOAFQDKNHLVMEYQPG 180
 DB 121 DVYAMKIMKKKALLAEOQVSPFEEERNILSRSTSPMIPOLOAFQDKNHLVMEYQPG 180
 QY 181 DLSLNRYEDQDENLQFYLAELILAVSHVLMGVHVDIKPENILVDRTHIKLVDF 240
 DB 181 DLSLNRYEDQDENLQFYLAELILAVSHVLMGVHVDIKPENILVDRTHIKLVDF 240
 QY 241 GSAAKNSNMKNVNAKLPIDPDYAPVLTVMNGDGKGTGLDCDMWSVGVIAYEMTYGR 300
 DB 241 GSAAKNSNMKNVNAKLPIDPDYAPVLTVMNGDGKGTGLDCDMWSVGVIAYEMTYGR 300
 QY 301 SPFAEGTSARTFNINMFORFLKPPDDPKVSSDFDLIOSLICGQKRLKEGICHPFF 360
 DB 301 SPFAEGTSARTFNINMFORFLKPPDDPKVSSDFDLIOSLICGQKRLKEGICHPFF 360
 QY 300 TPFTBGTSARTFNINMFORFLKPPDDPKVSSDFDLIOSLICGQKRLKEGICHPFF 359
 DB 300 TPFTBGTSARTFNINMFORFLKPPDDPKVSSDFDLIOSLICGQKRLKEGICHPFF 359
 QY 361 SKIDMNNIRNSPPFVPTLKSDDDTSNFDEPKNSWVSSPCQSPSGFSGEELPFVGF 420
 DB 361 SKIDMNNIRNSPPFVPTLKSDDDTSNFDEPKNSWVSSPCQSPSGFSGEELPFVGF 420
 QY 360 ARTDMNNIRNSPPFVPTLKSDDDTSNFDEPKNSWVSSPCQSPSGFSGEELPFVGF 419
 DB 360 ARTDMNNIRNSPPFVPTLKSDDDTSNFDEPKNSWVSSPCQSPSGFSGEELPFVGF 419
 QY 421 YSKALGILGRSESVSGLDSPAKTSMEKLLIKSKELQSDQCHXK 468
 DB 421 YSKALGILGRSESVSGLDSPAKTSMEKLLIKSKELQSDQCHXK 467
 RESULT 10
 088527
 ID 088527 PRELIMINARY; PRT; 448 AA.
 AC 088527;
 DT 01-NOV-1998 (Tremblrel. 08, Created)
 DT 01-NOV-1998 (Tremblrel. 08, Last sequence update)
 DT 01-MAR-2004 (Tremblrel. 26, Last annotation update)
 DE Cyttron-K kinase (Fragment).
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 NCBI_Taxid=10116;
 RN
 RP SEQUENCE FROM N.A.
 RA Medaule F., Eda M., Watanabe N., Fujisawa K., Matsunaka T., Bito H.,
 RA Ichizaki T., Narumiya S.;
 RT "Role of Cyttron kinase as a target of the small GTPase Rho in
 RT cytolysis";
 RL Nature 0:0-0(1998).
 CC
 CC SIMILARITY: Belongs to the Ser/Thr protein kinase family.
 DR EMBL; AF070065; AAC27932.1; -.
 DR HSSP; P31751; IMRV.
 DR GO; GO:0005524; F:ATP binding; IEA.
 DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
 DR GO; GO:0016740; F:transferase activity; IEA.
 DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
 DR InterPro; IPR011009; Kinase_Like.
 DR InterPro; IPR00961; Kinase_C.
 DR InterPro; IPR000719; Prot_kinase.
 DR InterPro; IPR002290; Ser_Thr_kinase.
 DR InterPro; IPR008271; Ser_Thr_kinase.
 DR Pfam; PF00069; Kinase_1.
 DR Pfam; PF00433; Kinase_C_1.
 DR ProDom; PD000001; Prot_kinase_1.

DR SMART; SM00220; S_TK_X; 1.
 DR SMART; SM00133; S_TK_X; 1.
 DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.
 DR PROSITE; PS00101; PROTEIN KINASE DOM; 1.
 DR PROSITE; PS00108; PROTEIN KINASE ST; 1.
 KM ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
 FT NON TER 448 448
 SQ SEQUENCE 448 AA; 50650 MW; 548CA5A8EA46394 CRC64;
 Query Match 20.6%; Score 2160.5; DB 2; Length 448;
 Best Local Similarity 91.1%; Pred. No. 6.3e-55;
 Matches 409; Conservative 20; Mismatches 19; Indels 1; Gaps 1;
 QY 1 MLKRYGARNPLDGAAPIASRASRLNLFQGGKPPMTQOQMSPLSREGILDALFVLE 60
 DB 1 MLKRYGARNPLDGAAPIASRASRLNLFQGGKPPMTQOQMSPLSREGILDALFVLE 60
 QY 61 ECSOPALMKIKHVSNFVKYSDTIAEIOEPSAKDFEVSILVCGHFAEYQVVERKATG 120
 DB 61 ECSOPALMKIKHVSNFVKYSDTIAEIOEPSAKDFEVSILVCGHFAEYQVVERKATG 120
 QY 121 DIYAMKVMKKKALLAEOQVSPFEEERNILSRSTSPMIPOLOAFQDKNHLVMEYQPG 180
 DB 121 DIYAMKVMKKKALLAEOQVSPFEEERNILSRSTSPMIPOLOAFQDKNHLVMEYQPG 180
 QY 121 DVYAMKIMKKKALLAEOQVSPFEEERNILSRSTSPMIPOLOAFQDKNHLVMEYQPG 180
 DB 121 DVYAMKIMKKKALLAEOQVSPFEEERNILSRSTSPMIPOLOAFQDKNHLVMEYQPG 180
 QY 181 DLSLNRYEDQDENLQFYLAELILAVSHVLMGVHVDIKPENILVDRTHIKLVDF 240
 DB 181 DLSLNRYEDQDENLQFYLAELILAVSHVLMGVHVDIKPENILVDRTHIKLVDF 240
 QY 241 GSAAKNSNMKNVNAKLPIDPDYAPVLTVMNGDGKGTGLDCDMWSVGVIAYEMTYGR 300
 DB 241 GSAAKNSNMKNVNAKLPIDPDYAPVLTVMNGDGKGTGLDCDMWSVGVIAYEMTYGR 300
 QY 241 GSAAKNSNMKNVNAKLPIDPDYAPVLTVMNGDGKGTGLDCDMWSVGVIAYEMTYGR 299
 DB 241 GSAAKNSNMKNVNAKLPIDPDYAPVLTVMNGDGKGTGLDCDMWSVGVIAYEMTYGR 299
 QY 301 SPFAEGTSARTFNINMFORFLKPPDDPKVSSDFDLIOSLICGQKRLKEGICHPFF 360
 DB 301 SPFAEGTSARTFNINMFORFLKPPDDPKVSSDFDLIOSLICGQKRLKEGICHPFF 360
 QY 300 TPFTBGTSARTFNINMFORFLKPPDDPKVSSDFDLIOSLICGQKRLKEGICHPFF 359
 DB 300 TPFTBGTSARTFNINMFORFLKPPDDPKVSSDFDLIOSLICGQKRLKEGICHPFF 359
 QY 361 SKIDMNNIRNSPPFVPTLKSDDDTSNFDEPKNSWVSSPCQSPSGFSGEELPFVGF 420
 DB 361 SKIDMNNIRNSPPFVPTLKSDDDTSNFDEPKNSWVSSPCQSPSGFSGEELPFVGF 420
 QY 360 ARTDMNNIRNSPPFVPTLKSDDDTSNFDEPKNSWVSSPCQSPSGFSGEELPFVGF 419
 DB 360 ARTDMNNIRNSPPFVPTLKSDDDTSNFDEPKNSWVSSPCQSPSGFSGEELPFVGF 419
 QY 421 YSKALGILGRSESVSGLDSPAKTSMEK 449
 DB 421 YSKALGILGRSESVSGLDSPAKTSMEK 448
 RESULT 11
 09VTV8
 ID 09VTV8 PRELIMINARY; PRT; 1854 AA.
 AC 09VTV8;
 DT 01-MAY-2000 (Tremblrel. 13, Created)
 DT 01-OCT-2002 (Tremblrel. 22, Last sequence update)
 DT 01-MAR-2004 (Tremblrel. 26, Last annotation update)
 DE CG10522-PA.
 GN ORFNames=CG10522;
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 NCBI_Taxid=7227;
 RN
 RP SEQUENCE FROM N.A.
 RA MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;
 RA Adams M.D., Celisner S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.C., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.G., Rogers Y.H., Blaziel R.G., Champs M., Pfeiffer B.D.,
 RA Man K.H., Doyle C., Baxter B.G., Helt G., Nelson C.R., Gabor G.L.,
 RA April J.F., Agbayani A., An H.J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.W., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benoit P.V., Berman B.P., Bhandari D., Bolashkov S.,
 RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,

RA Burtie K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cavley S., Dahler C., Davenport L.B., Davies P.,
RA de Padlos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.U., Wei M.H., Ibegwan C.,
RA Jajali M., Kalush F., Karpen G.H., Ke Z., Kennison U.A., Ketchum K.A.,
RA Kimmel B.E., Kodira A.A., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Laeko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Mekulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nuskeken D.R., Paclob J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H.,
RA Shue B.C., Siden-Klamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C.D., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodgerger, Worley K.C., Wu D., Yang S., Yao Q.A., Ye J.,
RA Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster";
RL Science 287:2185-2195(2000).
[2]
RP SEQUENCE FROM N.A.
RX MEDLINE=22426065; PubMed=12537568;
RA Celniker S.E., Wheeler D.A., Krommiller B., Carlson J.W., Halpern A.,
RA Patel S., Adams M., Champe M., Dugan S.P., Frise E., Hodgson A.,
RA George R.A., Hoskins R.A., Laverly T., Muzny D.M., Nelson C.R.,
RA Paclob J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,
RA Svirskas R., Taber P.E., Man K., Stapleton M., Sutton G.G., Venter C.,
RA Weinstein G., Scherer S.E., Myers E.W., Gibbs R.A., Rubin G.M.;
RT "Finishing a whole-genome shotgun: Release 3 of the Drosophila
melanogaster euchromatic genome sequence";
RL Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).
[3]
RP SEQUENCE FROM N.A.
RX MEDLINE=22426070; PubMed=12537573;
RA Kaminler J.S., Bergman C.M., Krommiller B., Carlson J., Svirskas R.,
RA Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,
RA Ashburner M., Celniker S.E.;
RT "The transposable elements of the Drosophila melanogaster euchromatin:
a genomics perspective";
RL Genome Biol. 3:RESEARCH0084-RESEARCH0084(2002).
[4]
RP SEQUENCE FROM N.A.
RX MEDLINE=22426069; PubMed=12537572;
RA Miera S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
RA Hradecky P., Huang Y., Kaminler J.S., Millburn G.H., Prochnik S.E.,
RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu U., Betman B.P.,
RA Bellocourt B.R., Celniker S.E., de Grey A.D., Drysdale R.A.,
RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.O.,
RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
RA Lewis S.E.;
RT "Annotation of the Drosophila melanogaster euchromatic genome: a
systematic review";
RL Genome Biol. 3:RESEARCH0083-RESEARCH0083(2002).
[5]
RP SEQUENCE FROM N.A.
RG FlyBase;
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
[6]
RP SEQUENCE FROM N.A.
RG FlyBase;
RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: Belongs to the Ser/Thr protein kinase family.
DR EMBL; AB003541; AAF49907.2; -.
DR HSSP; P31751; IMRY.

DR Intact; Q9VY78; -.
DR FlyBase; FBgn0036295; CG10522.
DR GO; GO:0000910; P:cytokinesis; IMP.
DR InterPro; IPR001180; Citron.
DR InterPro; IPR002219; DAG_PE-bind.
DR InterPro; IPR011009; Kinase like.
DR InterPro; IPR000961; Kinase C.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR002290; Ser_thr_kinase.
DR InterPro; IPR008271; Ser_thr_pkin_AS.
DR Pfam; PF00780; CNH; 1.
DR Pfam; PF00069; Kinase; 1.
DR Pfam; PF00433; Kinase_C; 1.
DR SMART; SM00109; Cl_1_C; 1.
DR SMART; SM00220; S_TKc; 1.
DR SMART; SM00133; S_TK_X; 1.
DR PROSITE; PSS0081; DAG_PE_BIND_DOM_2; 1.
DR PROSITE; PSS0107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PSS0011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PSS0108; PROTEIN_KINASE_ST; 1.
DR ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 1854 AA; 211323 MW; 93E5B69FB63DEBD CRC64;

Query Match 17.4%; Score 1822; DB 2; Length 1854;

Best Local Similarity 26.7%; Pred. No. 1.1e-44;

Matches 556; Conservative 406; Mismatches 646; Indels 474; Gaps 72;

QY 18 EPIASPASRLNLFQGGP-----PMTQOQMSPLS-----RAG 50
DB 6 EPIASVARTANLNILKKGAGVCAKPAAGSAGSGIPASTRSIYVSTTSAVAEACRGG 65
QY 51 IIDLAVLFEBSQPALMK-IRKVSNFKVRSPTTIAELOPSADQFEVRSVGGGHA 109
DB 66 IIDLAFCLVNECDKDTLTKRDRNIAEFVKKFRPIIETKRLRNADDFIKTLIGGYRG 125
QY 110 EYGVAREKATGDIYAKWKKKALLAEOVSFEEERNILTSRSTSPWIOYAFODKH 169
DB 126 NVHLVVERQNDIYAKKIKKSVVTSQ-----VKESRDWSIRNSMLNLNLOAFQNDN 181
QY 170 LYLWVEYQPGDILLSLNRYEDOLDENLIQFYLAELILAVSHVHMGVYHDIKPNILV 229
DB 182 LYLWVEYMGDGLSLMSRH-GPFEDDLARFYLAELTVLHNLTHMGVYHDIKPNILV 240
QY 230 DRTGHTKLVDFSGAAYKNSKRVNAKLPITGPDYNAEVLVWNGD-GKITYGLDCDMW 287
DB 241 DRTGHTKLVDFSGAAYKNSKRVNAKLPITGPDYNAEVLVWNGD-GKITYGLDCDMW 300
QY 288 SVGVIVYEMTYGSGPPAEGTSARTENNIM-----NFGRIKPPDDPKYSODFLDIQSL 341
DB 301 SWGIIIGELICETTPHEDNVHETYSKILSHCESHLKELISPPADLKVSVNRNLESL 360
QY 342 LCGQERLKEBGLCHPFFSKIDMNNIRNSPPFVFTLKSDDTSNFD-----PERN 394
DB 361 VYMSRRLSYERIKNHPFSEIPWMSIRSQVPIITVYASDDTSNFDGIRHKTRREG 420
QY 395 SWVSSPQCLSPSGFSGEELPYVGSYKALGILGRSESVSGLDSPATSSMEKLLIK 454
DB 421 VAKKSILTTMKNSNDPFGKDLPIFGYFVH-----MEKSAISATDEKLOEK 466
QY 455 SKK-----LQDSQCKHKMQEMTRHL---RVSEVAVYLSQKEVELKASRTQSLLEQL 507
DB 467 LKELLOKLTRENEISMLKODLLRAOOSLRKDNKSQVADAMEIK-KLQOIIKEKTM 524
QY 508 ATYITESSLSKSLK-QARMEVSQEDDKALQLLHDIRESRKLQETK---EOEYQAVNE 563
DB 525 E-LITCKQIKTLQSGAKIDEEMSKKATITDILRLRKQTEBEKINSQRYEQLAD 582
QY 564 MRLMNNQLBEDLVASARRSDLYSESLRESRLAAEEFKKATPCQK---KLRAKDGKPE 620
DB 583 KK-----QELASTIQKLDAREL-----EFNAKFECKHLSMKLQYKDWLQOI 625
QY 621 VGEYATLEKINAEQOLKIOLEKLEKAVKASTAEATELLQNTROAKE---PARELEKYO 677

```

Db 626 KEONLKSEPTNHEBOQRQMAELIYEQ-----KLTDLRKVKVRSDQDTRRMTWEIKEIR 676
Qy 678 NREDSSEICIRKALVEA-----EERRHLENKYKLETWERRENNIKDDIQTYSQOIQ 729
Db 677 TELDESISSSKSTOBAKUNATERNIEILRLNEIASNNELIAEKKVLETQLKNNETQ 736
Qy 730 QMADKILELEKHEAQSAGHLEVLKQKQHYE-----EKKVLDNOIKD-LADKE 782
Db 737 EVRAECHRLEHLOLAECRCQLAESSLATQVSPYETAPGSLTELNALIEDOLRADLAAKE 796
Qy 783 T-----LENNMORHEEAHEKGIUSEQA-----MINAMDSKIRS 818
Db 797 SENHQGRADLOLQVLVTKLEOMLERFNGQSLSPTKSHSRQOEBGTVDMLERONEKLED 856
Qy 819 -----LEQRIYSEANKLANSSILFTORNNKAOEEMISE---LRQOKFYLETQAGKLEA 870
Db 857 KLAAYRQMIYERQOAR--TANLSIM--KVEKQLEALSEKKLARRELTEDRTIKKQVN 912
Qy -871 ONRLEBOLEKISHODSDKRLLELETRL-----REVSLEH-----EBOQLKLRQLTE 920
Db 913 ASDAQRMLTKTSQEBTRQRESRIEELKQELAAAKRDVLKEHROWEKAQEERMKCKSEIIE 972
Qy 921 LQSLQRESQUTALQARALAESOLRQAKTELEBTTAEABEIOALTANHDEIQRKDA 980
Db 973 HLANVHLEQOETELR-----QKRLQISRFPGVTILEQKNTTIRELOEERKSKRANKD- 1024
Qy 981 LRNSCTVITDLEQALNOLTEDNAELNNQNFYLSKQDLEASGANDIYOLRSEVHLRREI 1040
Db 1025 ---SCIV---LQKELKQLTID-----NF---QRLKVCAGSITDS-----QL 1054
Qy 1041 TEREMOLTSQOKOTMEALKTTCTMLEBOVMDBALNDELLEKEROMEANRSVLGDEKSOFE 1100
Db 1055 TEVETMLKSEBERKRSOKS-----QDITLHEKIRENDDQ----- 1088
Qy 1101 CRVELORMJLDTKOSRABADQ---ITESHQVVELVKEHKAELILQOALXEOKLKAE 1157
Db 1089 --LIDLKQLTIVSEKRLAEQRAQVLASETIDELRLNKEKQKCVAAQDQVLBO----- 1141
Qy 1158 SLSDKLDNLEKGMALB-MNARSLOQKLETERELKORILLEBOAKLQOQMDLQKNHIFL- 1215
Db 1142 --TALFATQORARILBDQNNANYEAQTPADSREWVS--LKENANILISLFIKKEVEGNLQ 1198
Qy 1216 --TQGLQALDRADLLKLTENSLEYOLENIOVLVSHKVKMEGTSQOTKLIDFLQAM- 1272
Db 1199 AEIRGLESA--QAN-LHAELIDSLDTLAEKQFYVORDIKSNATLAQHKKIDVLIQKVE 1255
Qy 1273 DQPAKKKKV-----PQVMEKALKKEKARCAELEALQKT 1309
Db 1256 DLSAKKKKTLADKLPGSSHTNKENVSPNDVSSITLYRALKEKLRREQMNSLKEQLAQL 1315
Qy 1310 R--TELRSABEAAHRAKATDHPHPSTPATARQOIAMSAIVRSPHOPSAMSLAPSSRR 1367
Db 1316 NGRTILASPRKSAVANDSDAP-----KORPVSIALPRSPKQOOP--LKRITTSQV 1365
Qy 1368 KESSTPEEFSSRLKERMHNIPIHPFNVL--NNRATKCAVCLDTVHFGRQASKCLEQV 1424
Db 1366 ELKTTAKETPVTTIENQAH--HRFELALQSKYDAVAVCVCEKAVVAGSPFMCKECKD 1422
Qy 1425 MCHPKCSTCLPATGGLPAEYVITHTEAFCRKQMSPGIQTGEPSSSLHLEGMVAPRANK 1464
Db 1423 VTHRKCSSNVOSHG-----STKPTA----- 1443
Qy 1485 RGQOGMDRKVYVLBGSKVLIYDNEAREAGQRPVEEFELCLPDGVSITHGAVGASELANTA 1544
Db 1444 -----PSAD----- 1447
Qy 1545 KADVPYLLKMHSHPTTCWPGRTLLYLLAPSPDKQRWYALAESV--VAGGRVSRKABADA 1603
Db 1448 --DLSSISQVSS-----LTLDSYDVAGGTTSS-----GG 1473
Qy 1604 KLGNSLKLKEGDD-----RLDMNCTLPFSD-QVIVYGTBEGLYALANVLKNSLTHVIGIG 1657
Db 1474 EYIGSLVYSSDGAEDQARKEIEVCAFEVAAEQIILLGCGNTGLAYHLDQRLVHISGLE 1533

```

```

Qy 1658 AVFOIYIITKLEKLMT--AGEERALCLVDYKVK--QSLAQSHLPAPQDIS-----PN 1707
Db 1534 SVSCMSICKRLAKAIMVGTQBEKLYOC--DYROLESRCQSSSSCHKPLRTSALELPPAN 1591
Qy 1708 IFEAVKGOHLEGACKINGU--CICAMPKSVVILRYNENLSKXCIRKEIFSEPCGCIH 1765
Db 1592 RTPBEKMKVLISIEBAENALDSVAIAATSTIVLTKLDLKHMFKPVADLTAFTVTSIF 1651
Qy 1766 FTNYSILIGTNKFEYIDMKQYTLLEFLDKNDHSLAPVFAASNSFPVSIYQVNSAGORE 1825
Db 1652 FTRHSALVSDPKFEIDLNDVNAAEFVDSLQSKW-----ESTACQGPLTAVRIS-----RQ 1702
Qy 1826 EYLICFHEFGVFDVSYGRSRSTDDLKMSRLPLAFAYRPPYLFVTHFNSLEVIET----- 1879
Db 1703 EYLICFAEYGVFVDFECRSRPRYDLNWWYAPTGFVYRDPFLFISHYQSVQIVRLHRSFSK 1762
Qy 1880 -----QARSSAGRP--ARAVLDIPNRYLGPAL-SSGAYIL 1912
Db 1763 EMASGDNASENSESPELQRYVL---PHYMSTLLANSQDVNL 1800

RESULT 12
ID 054874 PRELIMINARY; PRT; 1732 AA.
AC 054874;
DT 01-JUN-1998 (TREMblrel. 06, Created)
DT 01-JUN-1998 (TREMblrel. 06, Last sequence update)
DT 01-MAR-2004 (TREMblrel. 26, Last annotation update)
DE Myronic dyctrophy kinase-related Cdc42-binding kinase.
GN Name=NRCK;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=98078670; PubMed=9418861;
RA Leung T., Chen X.-Q., Zan I., Maneer E., Lim L.;
RT "Myronic dyctrophy kinase-related Cdc42-binding kinase acts as a
RT Cdc42 effector in promoting cytoskeletal reorganization."
RL Mol. Cell. Biol. 18:130-140 (1998).
CC -I- SIMILARITY: Belongs to the Ser/Thr protein kinase family.
CC -I- SIMILARITY: Contains 1 PH domain.
CC EMBL: AF021935; AAC02941.1; -.
DR PIR: T14039; T14039.
DR HSBP; P31751; IMRY.
DR GO: GO:0005524; P:ATP binding; IEA.
DR GO: GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR GO: GO:0005083; F:small GTPase regulatory/interacting protein. ., IEA.
DR GO: GO:0016740; F:transferase activity; IEA.
DR GO: GO:0007242; P:intracellular signaling cascade; IEA.
DR GO: GO:0006468; P:protein amino acid phosphorylation; IEA.
DR InterPro: IPR002119; DAG_PE-bind.
DR InterPro: IPR001180; Citron.
DR InterPro: IPR002219; DAG_PE-bind.
DR InterPro: IPR011009; kinase like.
DR InterPro: IPR001304; lectin_C.
DR InterPro: IPR000095; PAKbox/rhoindng.
DR InterPro: IPR001849; PH.
DR InterPro: IPR000961; Pkinase_C.
DR InterPro: IPR000719; Prot kinase.
DR InterPro: IPR002290; Ser Thr pkinase.
DR InterPro: IPR008271; Ser_thr_pkin_5s.
DR Pfam; PF00130; CL_1.1.
DR Pfam; PF00780; CNH; 1.
DR Pfam; PF00169; PH; 1.
DR Pfam; PF00069; Pkinase; 1.
DR Pfam; PF00433; Pkinase_C; 1.
DR PRINTS; PR00008; DAGPDOMAIN.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00109; CL; 1.
DR SMART; SM00036; CNH; 1.

```

DR SMART; SM00285; PBD; 1.
 DR SMART; SM00233; PH; 1.
 DR SMART; SM00220; S.TK; 1.
 DR SMART; SM00133; S.TK X; 1.
 DR PROSITE; PS50108; CRIB; 1.
 DR PROSITE; PS00615; C TYPE LECTIN 1; UNKNOWN_1.
 DR PROSITE; PS00479; DAG_PE_BIND_DOM_1; 1.
 DR PROSITE; PS50081; DAG_PE_BIND_DOM_2; 1.
 DR PROSITE; PS50003; PH DOMAIN; 1.
 DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.
 DR PROSITE; PS50011; PROTEIN KINASE DOM; 1.
 DR PROSITE; PS00108; PROTEIN KINASE ST; 1.
 DR ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
 KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
 SQ SEQUENCE 1732 AA; 197062 MW; 65f62f7872ACD3B CRC64;

Query Match 14.5%; Score 1524.5; DB 2; Length 1732;
 Best Local Similarity 24.8%; Pred. No. 3.5e-36;
 Matches 523; Conservative 335; Mismatches 667; Indels 599; Gaps 76;

QY 36 PFMTQGMSPLSREGIIDLALFVLEECSPALMKIKVSNFVKYSDTIAELQELQPSAK 95
 DB 18 PAQNGQC--FSVETLADILICLYDECNNSPLRKKNILEYEWAKPFTSKVMRLHRE 75
 QY 96 DFEVRLVGCHFAFVQVREKATGDIYAMKVMKKKLLAEOQVFPFEERNILSRSTP 155
 DB 76 DFEILKVIIGCAFGEVAVVKLKADKVFAMKILNKMEMLKKAETACREERDVLVNGDSK 135
 QY 156 WIPLOVAFODKNLIVMEYQPGDLSLNLRYEDQDENTLIQVLAELIHAVSHLM 215
 DB 136 WITTLHAFODDNLLVMDYVGGDLTLTLTKEDLPEEMAFYLAENVIALDSYHQL 195
 QY 216 GYVHRDIKPNILVDRTHI KLVDFGSAKNSKRVNAKLPITGPDYMAPEVLTVANGD 275
 DB 196 HYVHRDIKPNILVMDNGHIRLADFGSCLKLMEDGVQSSVAVGTPTYISPILLQAME-D 254
 QY 276 GKGTGGLDCPMWSGVVAIEMITIGRSPABSTAKRTNININPRLKLPD-PKVSDF 334
 DB 255 GKRGYGECPMWSGVVCMYEMLYGETPYAESLIVETGKIMNHKRPQFPQVTVSENA 314
 QY 335 LDILOSLGQKELKEFGL---CCHPFFSKIDNNINRSPRPPTKSDSDTSNDEP 391
 DB 315 KDLIRLIRICREHRLGONGIEDFKHPPFGSIDMDNINRNCARPTIPEVSEPTDINSFVD 374
 QY 392 E---KNSWSSSPQCLSPSGSGELPFVGFVSYSKALGILGRS-----ESV 434
 DB 375 DDCLKNSGEMTPP---THTAFSGHHLPFVGTYSSTCLSDRSCLRTAGPTSLDLVNV 431
 QY 435 VSGIDSPAKTSSMKLLIISKELQDSODCKHKEOEMTRILHRVSE---VEAV-LSQK 489
 DB 432 QRTLDNNILATEAYERRI-----KRLQEKLELTRKLOESTQVVALQYSTV 477
 QY 490 EVELKASQTSRLLEOPLATYTCSSIKSLQEAQMEVSOEDKALQLLHIDINEQSKL 549
 DB 478 DGPITLTA-----KDL---EIKSLKEITEKRLQVAVEN-----507
 QY 550 QEIKGEYOQAQVEEMRLMNOLEEDLVASRRSDLYSELSERLAAEFGRKATQCOHK 609
 DB 508 -----HLEQQLER---ANSVRELDLDAFRQIKAFEQIKTLQOEELINKELVQVSR 557
 QY 610 LIAKADQCKPEVGYALGKINAEOQLIQLOELKEKAVASTEATELLONITQAEKRA 665
 DB 558 L---KNSK-----ELKDAHCOCKLAMQEFMEINERLTETLHQOKLARHVDKKEEV 607
 QY 670 ERELKEQNEDSEGIKRLKVEAEERRHSLENVKVLLETMERENRLLKODIQKSOQIQ 729
 DB 608 DLVWQK-----AESIRQELRRAERAKELVETTEALIAASCKRKLREGRSRYISQLE 660
 QY 730 QMADKILIELEKREAVSAQHLVHLKQKQHYEKKIKVLDNO-----IKQDLADKETL 784
 DB 661 N-----ELE-----GLKQKQISYSPGICSEIHQOEITKKTDLKKSIF 699
 QY 785 ENNMQRHEEAEHKGKLLSQKAMINAMDSKIRSLQRIEVLSTANKLANSSIFTORNM 844

DB 700 -----YEEELSRREGIHA-----SEIKNLKKEELHND-SEGOALANKKEIMV----- 738
 QY 845 KAEQEMISBLRQCKFYLETQAGKLEAQNRKLEBOLEKISHQSDXKRLLELTRLEVS 904
 DB 739 -----LKKDLEK-----TR-RESQ 751
 QY 905 LEHEBQKLEIKRQTLQSLQREBSQTLTQAARALBSQLRQAKTELEETTAABEEI 964
 DB 752 SEREEFENEKQO-----YEREVL-----771
 QY 965 QALTARDEIQRFADALRNSCTVITLBEQNLQTEDMELNQNQNFYLSKQLDASGAND 1024
 DB 772 -----LTEENKTL-----780
 QY 1025 EIVQLSSEVDHLREITERMQTLQKQTEALKTTCTMLEBOVMDLALNDELKERQ 1084
 DB 781 -----SELD-----KLTSS---LYESLIRNQHLEEVKDLA---DKKSVAH 816
 QY 1085 WEAMRSVLDGKQSFECRVRELQRMIDTEKQSRARADQRTESQVVELAVKEHKAEILA 1144
 DB 817 WEA-----QITEIILQWVSEKDAAG-----YIQAALA 842
 QY 1145 LQOALKEOKLKAEISLQKNDLE---KKHMLEMNAK-SIQOKLETREBELQRLLEQAK 1200
 DB 843 SKATTEBELALRNSLQTRATDMKRRKRAKLDMSARLEQSLADAEIRAKQALQEBLNK 902
 QY 1201 LQOQMDLQKKNHIFRLTQGLQELADRADLKTERSDLEIYQLENIQVLSHEKVMGTSIQ 1260
 DB 903 VK-----ASNITTECKLKQSEKKNELSEIRQLIKDTEELSRKQVEH 946
 QY 1261 QTKLIDFLQAKMDQPAKKKVPLOVNEKLAKLEKARCALEBQAKRIEIRAREEA 1320
 DB 947 RDSQSEFL-AFLMTPT-----DALDQEE-RS-----970
 QY 1321 AHRKATDPHPS--TPATARQIAMSALVRSPEHQPSAMSLAPPSRRKESSTPEERSR 1379
 DB 971 -----PSTPQKGRRLDSAPL--PVNTPTL-----RKQCPASAGRPK 1009
 QY 1380 LKERMENHNIPIRPNVGLNMRATKCAVCLD-TVHFGQASKLEQVQVCHPKSTCLPATC 1438
 DB 1010 RK-----THQFVFSFPAFTKCHQCTSLMVLIRQGCSEVGFSCCHITCVNAKPTTC 1062
 QY 1439 GLPAEYATHTFAFCQDKNNSPGLQTER-----SSSLHIGMMKVPNNNRG-QQGV 1490
 DB 1063 PVPE-----QTKGPIGIDPQKGVGTAYGHWIRPK--PAGVKGW 1101
 QY 1491 DRKYIVLESGKVLIVNEAREAGO-RVBEFEELCPRGDVSIGAVGASLANTAKADVP 1549
 DB 1102 QRLAVVCPKFLVYDIABGKASQPSYSIQVIDKDEPFSV--SSVLASDVHASKRIDP 1160
 QY 1550 YILKME---SHPTTCWPGRTLYLLAPSPDPKQWVTLAESVAVAGRVSRKAEADAKL 1605
 DB 1161 CIFRYTASQSLASPSDK-----SLTMLADSETSESKNVGVLSL---HKVLKKNKFRDRSV 1213
 QY 1606 LQNSLKLQGDRLDMNCTLP-----PSDQVYVLTETEGYALANVLKNSLTHVPG 1655
 DB 1214 ---YVPKEAYD-----STLPILKTTQAAIIDHERVALCNEBGLFVAVHTQKELLRVGD 1264
 QY 1656 IGAVFOIYIILKLEKLMITAGEERALCLVDYKVKQSLAQSHLPAQDISPNIFEAVKYG 1715
 DB 1265 NKKIHQELLPQOLAVAVISGRNRHVALPFGMSLYDRETFYKLA-----TKGC 1314
 QY 1716 HLFAGKTIENGL--CICAMPKRVVILRYVENLSKYCIRKEIETSEPCSIHFTNYS--I 1771
 DB 1315 QTLAAGVRRHAGALSCLVANKRQVLCBELPQSTRHKKFEIQV--PCNVQMAIFSEHL 1372
 QY 1772 LIGTNKFEYIDMKQYTL-----EFLDKDHSILAPVPAASNSPVSIVQVNSAQRE 1825
 DB 1373 CVG---FQSGFLRYPLNGRSGPCNMILHSDHTLA--FITTHQPMALCAVBSIN---K 1421
 QY 1826 EYLLCFHEBGFVFDVSGRGRSRTDLDKMSRLPLFAVAREPLFTPHFNSLVIELOARSSA 1885
 DB 1422 EYLLCFSSIGIYTDCCGRSRQOELMMPANPSSCCTNAPYLSIYSENAVVIDPVNSMWEI 1481

QY 1886 GTPARAVIADIPNRYLGPATSSGAIYLAASSYQDYLRYCCKGNLVKES-----GTEHHRG 1940
 Db 1482 QT-----LPLKQVR-----PLNTGSLINLLG--LETIRILRYFNKRAEGDELVLVPETSDNSR 1531
 QY 1941 PSTSRSSPNKR-----GPPTYNE-----HI 1960
 Db 1532 KQMKRNINNNKRRYFRVPEERMOQRREMLDPEKRNKLINPTNFHIAHMGFGDGIQT 1591
 QY 1961 TKRVASSPAPPEG-----PSHREPESTPHRYRGRTYL--RRDKSPGRPLEREKS 2008
 Db 1592 LKQJPMNRPPESESTVFSGSVSIPTIKSRPEGRSMASGLSARSSAQNGSLAKTEFS 1651
 QY 2009 PGRILSTRERKSPARLFEDESRGLPAGAV 2038
 Db 1652 GGS-YNTKROPMPG-----PSEGLSSGCV 1675

RESULT 13
 ID 086XX3 PRELIMINARY; PRT; 1719 AA.
 AC 086XX3;
 DT 01-JUN-2003 (Tremblrel. 24, Created)
 DT 01-JUN-2003 (Tremblrel. 24, Last sequence update)
 DT 01-MAR-2004 (Tremblrel. 26, Last annotation update)
 DE CD042 binding protein kinase alpha (DMPK-like).
 GN Name=CD042BPA;
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Colon;
 RA Wilkinson S., Marshall C.J.;
 RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: Belongs to the Ser/Thr protein kinase family.
 CC -1- SIMILARITY: Contains 1 PH domain.
 DR EMBL; AJ518975; CAD57745.1; -.
 DR HSSP; P31751; IMRY.
 DR GO; GO:0016030; C:membrane; IEA.
 DR GO; GO:0005524; F:ATP binding; IEA.
 DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
 DR GO; GO:0005083; F:small GTPase regulatory/interacting protein. .; IEA.
 DR GO; GO:0016740; F:transferase activity; IEA.
 DR GO; GO:0007242; P:intracellular signaling cascade; IEA.
 DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
 DR InterPro; IPR001180; Citron.
 DR InterPro; IPR002219; DAG_PE-bind.
 DR InterPro; IPR011009; Kinase_like.
 DR InterPro; IPR01304; Lectin_C.
 DR InterPro; IPR003345; M_repeat.
 DR InterPro; IPR000095; Pakbox/Rhoindng.
 DR InterPro; IPR001849; PH.
 DR InterPro; IPR000961; Kinase_C.
 DR InterPro; IPR000719; Prot_kinase.
 DR InterPro; IPR002290; Ser_thr_kinase.
 DR InterPro; IPR008271; Ser_thr_pkin_AS.
 DR Pfam; PF00130; Cl_1; 1.
 DR Pfam; PF00780; CNH; 1.
 DR Pfam; PF02370; M; 1.
 DR Pfam; PF00169; PH; 1.
 DR Pfam; PF00069; Pkinase; 1.
 DR Pfam; PF00433; Pkinase_C; 1.
 DR PRINTS; PR00006; DAGPEDOMAIN.
 DR ProDom; PD000001; Prot_kinase; 1.
 DR SMART; SM00109; Cl_1.
 DR SMART; SM00036; CNH; 1.
 DR SMART; SM00285; PBD; 1.
 DR SMART; SM00233; PH; 1.
 DR SMART; SM00220; S_TKC; 1.
 DR SMART; SM00133; S_TK_X; 1.
 DR PROSITE; PS50108; CRIB; 1.

DR PROSITE; PS00615; C_TYPE_LECTIN_1; UNKNOWN_1.
 DR PROSITE; PS00479; DAG_PE_BIND_DOM_1; 1.
 DR PROSITE; PS50081; DAG_PE_BIND_DOM_2; 1.
 DR PROSITE; PS50003; PH_DOMAIN; 1.
 DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
 DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
 DR KATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
 SQ SEQUENCE 1719 AA; 195949 MW; DB15FA8611D9871 CRC64;

Query Match 14.2%; Score 1489; DB 2; Length 1719;
 Query Local Similarity 24.1%; Pred. No. 3.6e-35;
 Matches 523; Conservative 330; Mismatches 628; Indels 666; Gaps 76;

QY 36 PFMTQGMSPLSRGRIIDALFVLEPESGOPALMKIKAVSNFVRKYSPTIAELQELQPSAK 95
 Db 18 PAQTNQGC--FSVETLIDILICLYDECNNSPLREKXNILEVEMAKPPTSXYKQMRLHRE 75
 QY 96 DFEVRSIVGCGHFAEVOVVEKATGDIYANKVMKKKLLAQEQVSPFEERNIISRSTP 155
 Db 76 DFEILKVIIGRGAFAEVAVVVKLNADKVFAMKILKWEMLKRAETACPREEDVLVNGDNK 135
 QY 156 WIPQLOVAFODKXNLYVMFQEPGDLISLNRVEDLDENLIQFYLAELILAVSHYM 215
 Db 136 WITTLHYAFQDDNNLYVMWDTYVGGDLTLTLSTKFEDELPEDMARFYLAEMVAIDSVHQL 195
 QY 216 GYVARDIKPENILVDRGTGHIKLVDFGSAKQNSKNMNAKPLIGTPDYMAPEVLTVNNGD 275
 Db 196 HYVARDIKPENIILMDNMGHIRLADFSGCLKMBEGTVQSSVAVGTPPYISIELIQANE-D 254
 QY 276 GKGTGYLDQWMSVGVIAIYEMITGRSPABGTSARTNNINNFQRLKFPDD-PRVSSDF 334
 Db 255 GKGRYGCECDWMSIGVCMYEMLYGETPFYASLVEITGYKIMNHKRRQFPQAVTVDSENA 314
 QY 335 LDIQLSLCQKREKREGL---CCHPEFSKIDNNINRNSPPVPVPLKSDDDPSNPEP 391
 Db 315 KDILRLRISCRERHLGNGIEDFKHPFSGIDMDNIRNCEAPYIPEVSSPTDTSNPDVD 374
 QY 392 E---KNSWVSSPQLSPSGSGEELPFVGSYKALIGRS-----ESV 434
 Db 375 DDCLAKSETPMP---THTAFSGHILPFVGFTTSSCVLSDRCLRTAGTSLDLVNV 431
 QY 435 VSGLDSPAKTSMEKKLLIKSKELQSDQCHKNGQEMTRILHRVSE---VEA----- 484
 Db 432 QRTLDNNLATEAVERRI-----KRLQEKELSRKLQSTQVQALQYSTV 477
 QY 485 ---VLSQKEVELKASEQSLLE--QDLATYITGSSLKSLSEQARNEVQEDDKALQL 539
 Db 478 DGPLTASKDELIK-----NIKEIETLRQVYESSHLEQOLEBAN-AVRQELDDAFROI 530
 QY 540 HDIREGSRKLOETKEQEVYQAVEEMRLMMNQLEEDIVSABRRSDLYESELRESRLAABF 599
 Db 531 KAYEKQIKTLQERE-----DLNKELVQASERLKNQSKELKDA----- 568
 QY 600 KRKATECOHKLAKAQGKPEVEGYALTEKINAEQOLKIOLEKLEKAVASTEATEEL 659
 Db 569 -----HCQRK-----LAMQEPMEINERLREIHLHQXKLA 597
 QY 660 QNIRQAKERARELEKIQNREDSSEGIRKKLVAEERHSLKENVKQLETMERENRKLKD 719
 Db 598 RHVADKEEVDLVWQKV-----ESLRQELRTERAKKELEVTEALAAEASDRKLRE 650
 QY 720 DITQSQIQIOWADKILIEKEKIREAOVSAQHLEVHLKQKQHYEEKIKYVLNDQ-----I 774
 Db 651 QSEHYSKQLEN-----ELE-----GLKQIQISIPVCGSIEHQOETTKL 689
 QY 775 KQDLADKETLENNQRRHEEAHEKGLISQOKAMINAMDSKISLEQRIYELSEANKLAA 834
 Db 690 KTDLEKKSIF-----YEE-----LSKREG-IHA-NEIKNLKKEIHD-SEGQQLAL 732
 QY 835 NSSLFTQRMKAQEMNISELRQOKFYLETQAGKLEAQRNKLQEOLEKISHQSDKRRLL 894
 Db 733 N-----KEIMI-----LKQLEK----- 745

Qy	895	ELERLEBVSLEHEQKLELROLTLEQLSIQEBSQUTALQARALLESQLRQAKTELE	954
Db	746	---TR-RSOSEREFESEFPQOYEREBVLTLEBKNTLS-----	781
Qy	955	ETTEAEBEIIOALTANRDEIORKFDALRNSCTVITDLEBQNLQTE--DNAELNNOFYL	1012
Db	782	-----	798
Qy	1013	SKOLDEASGANDEIVOLRSEVNDHLRREITEREMQLTOSQOMTEALKTTCTMLEQVNDLE	1072
Db	799	-----	808
Qy	1073	ALNDELLEKQWEMAMSVLGDSEKQSFECRYREIORMJDEKOSRARADQITRSQVVE	1132
Db	809	---DKSESVAHWE-----	836
Qy	1133	LAVENKHAKEIITALOQALKEQKILAKESLSDKINDLE--KKEHAMEMNAR-SLOQKLETER	1188
Db	837	-----YIQALASKMTBEELARNSLSOTRATIDMFWKRRARRAKIDMSARLEQALDAE	889
Qy	1189	ELKORLLEBQAKLOQMDLOKNHIFRLTQLOZALDRADLINTERSDLEIYOLENIQVLYS	1248
Db	890	-----	901
Qy	1249	HEKYMESTIQOQTKLIDFLQAKNDQPAKKKQVLPQVYELKLALEKAKARCAELEBQK	1308
Db	902	---KYASNIIT-ECKLKD-----	934
Qy	1309	TRIELSAREEAHARKATDH--PHSPATAPAOIAMSATVRSBEPHOSAMSILAPSS	1365
Db	935	TE-BIRS--EKGIERODSQHFLAFLNTPTDLDQF-----	982
Qy	1366	RRKESSTPEEFSRRLKERMHNNIPHRFVNGILMRYATKAVCLD--TVHGRQASKCLEQOV	1424
Db	983	RKKQCPGSGTGPFRKK-----	1035
Qy	1425	MCHEKCSCTLPATGCLPAEYATHTFEAFCDROMNSPGLOQTEP-----	1477
Db	1036	SCHTTCVNAKPAFTTQVPRPE-----	1076
Qy	1478	KVPENNKRG--QOGMDRKATVIEGSKVLTYDNDEARAGORPV-----	1526
Db	1077	RIPK--PAGVKKGORALATVCDKFLTYDIAEGKASOPSVIYQVINDMRBEBSV---	1130
Qy	1527	GDVSIHGAVGASELANATAKADVPLYLKME-----	1582
Db	1131	-----SSVLASDVIIHASRKDIPCFIPTYASQLSASNNKC-----	1180
Qy	1583	TALESVAVAGVRSEKAEADAKLGNLSLKIEGGDRL-----	1626
Db	1181	GVLSF-----	1222
Qy	1627	FSDQVVLVNGTEBGLYALVNLKNSLTHVPGIGAVPQIYIILKOLEKLMJAGERALCU---	1683
Db	1223	IDHEKIALAGNEBGLFVAVHTDDEIIRVDNKKIHQIBELIPNDQLAIVSGNNRHRRLPBM	1282
Qy	1684	-----VDVKKVKSLOASHLPAPORDISPNIPEAVKGCHEFGAGKLENG--LCICAMP	1734
Db	1283	SALDQRETDYFKLST-----	1332
Qy	1735	SKVYLIRYENLSKYCIKKEIETSEBSCIHFTNYSILIGTNKEYEIDMKOYTL-----	1789
Db	1323	QOVLCEYELFQSKTRHFRKEIQLPVPYVNMMAIFSQLCVQ-----	1378
Qy	1790	ELFDKNDHSLAPAVPAASNSFPVSIYQVNSAGREBYLLCFHEGFVNDYGRKSTD	1848
Db	1379	YSMHSNHTLS--FIHQPMDALCAVEIS-----	1431
Qy	1849	DLKMSRLDLAPAYRPPYFTVHFNLSYELIEIQARASAGTPRAYVYDIDNRPYLGPATSSG	1908
Db	1432	ELMFPANPSSCCTANAPVLSYSENAVDFDVNSMWTOT-----	1483

Oy	1909	AIYLAASSYQDKRYICCKGNLVKES-----GTEHHRRPSTRSSPNKR-----	1951
		:: ::	
Db	1484	SLNLG--LETRILTYFKMKABODELVPEPDSNKRQKQVNNKRRISFVPEEERM	1541
Oy	1952	-----GPETYNE-----HITKRVASSPAPPEG-----	1973
Db	1542	QQRREMLRDEPMRNKLI-SNPTNFNFIHWGPGDQIQLKLDPMNRPQESRTVFGSGSVSI	1601
Oy	1974	PSHREPESTPHRRREGTEL--RDKSPGPRLEREKSPPGILSTRRRSRPARLFEDSSRG	2031
		::	
Db	1602	PSIKRSPPEGRSWSASSGLSARSSAONGSLAKRFEGSGS--YSAKRQPMPS-----PSEG	1655
Oy	2032	RLPAGAV 2038	
Db	1656	SLSSGGM 1662	
RESULT 14			
O9ULUS			
ID	O9ULUS	PRELIMINARY;	PRT; 1760 AA.
AC	O9ULUS;		
DT	01-MAY-2000 (TREMBLrel, 13, Created)		
DT	01-OCT-2001 (TREMBLrel, 18, Last sequence update)		
DT	01-OCT-2003 (TREMBLrel, 25, Last annotation update)		
DE	KIAA1124 protein (Fragment).		
GN	Name=KIAA1124;		
OS	Homo sapiens (Human).		
OC	Eukaryota; Metazoa; Chordata; Cranilia; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.		
OX	NCBI_TaxID=9606;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	TISSUE=Brain;		
RX	MEDLINE=20039618; PubMed=10574461;		
RA	Hirosawa M., Nagase T., Ishikawa K., Kikuno R., Nomura N., Ohara O.;		
RT	"Characterization of cDNA clones selected by the Genemark analysis		
RT	from size-fractionated cDNA libraries from human brain."		
RL	DNA Ref. 6:329-336(1999).		
CC	-1- SIMILARITY: Belongs to the Ser/Thr protein kinase family.		
CC	-1- SIMILARITY: Contains 1 PH domain.		
DR	EMBL; AB032950; BAA86438.2; -.		
DR	HSSP; P31751; IMRY.		
DR	GeneW; HGNC:1738; CDQ42BP.		
DR	GO; GO:0005524; F:ATP binding; IEA.		
DR	GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.		
DR	GO; GO:0005083; F:small GTPase regulatory/interacting protein. . ; IEA.		
DR	GO; GO:0016740; F:transferase activity; IEA.		
DR	GO; GO:0007242; P:intracellular signaling cascade; IEA.		
DR	GO; GO:0006468; P:protein amino acid phosphorylation; IEA.		
DR	InterPro; IPR002119; Cltron.		
DR	InterPro; IPR002119; DAG_PE-bind.		
DR	InterPro; IPR011009; Kinase like.		
DR	InterPro; IPR000095; PAKbox/RhoBindg.		
DR	InterPro; IPR001849; PH.		
DR	InterPro; IPR000961; Pkinase_C.		
DR	InterPro; IPR000719; Prot_kinase.		
DR	InterPro; IPR002290; Ser_thr_pkinase.		
DR	InterPro; IPR008271; Ser_thr_pkin_AS.		
DR	Pfam; PF00130; Cl_1.1.		
DR	Pfam; PF00780; CNH; 1.		
DR	Pfam; PF00786; PBD; 1.		
DR	Pfam; PF00169; PH; 1.		
DR	Pfam; PF00069; Pkinase; 1.		
DR	Pfam; PF00433; Pkinase_C; 1.		
DR	PRINTS; PRK00008; DAGPEDOMAIN.		
DR	ProDom; PD000001; Prot_kinase; 1.		
DR	SMART; SM00109; Cl; 1.		
DR	SMART; SM00036; CNH; 1.		
DR	SMART; SM00285; PBD; 1.		
DR	SMART; SM00233; PH; 1.		
DR	SMART; SM00220; S_TKc; 1.		
DR	SMART; SM00133; S_TK_X; 1.		
DR	PROSITE; PS50108; CRIB; 1.		

DR PROSITE; PS00479; DAG_PE_BIND_DOM_1; 1.
 DR PROSITE; PSS0081; DAG_PE_BIND_DOM_2; 1.
 DR PROSITE; PSS0003; PH_DOMAIN; 1.
 DR PROSITE; PSS0107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE; PSS0011; PROTEIN_KINASE_DOM; 1.
 DR PROSITE; PSS0108; PROTEIN_KINASE_ST; 1.
 KM ATP-binding; kinase; Serine/threonine-protein kinase; Transferrase.
 FT NON_TER 1 1
 SQ SEQUENCE 1760 AA; 199208 MW; 3A1CA9A7A4BFSFA3 CRC64;
 Query Match 13.7%; Score 1440; DB 2; Length 1760;
 Best Local Similarity 24.4%; Pred. No. 9,4e-34;
 Matches 512; Conservative 336; Mismatches 668; Indels 584; Gaps 76;
 QY 44 SPUSREGILDALFVLFEECSQAPALMKI KHSNFPVKYSDTIAELOLPASAKDFEVS LV 103
 DB 72 SALSVETLIVLVCLTEYCSHSLRDKYVAEFLEMAKPFQVAKENQHLREDEILKVI 131
 QY 104 GCGHFAVOVREKATDTIYAMKMKKALLAQOVSPFEENITLSRSTSPWIPQLQYA 163
 DB 132 GRGAFGEVAVAKMNTIRIYAMKILINKEMIKRAETACFPERDVLVNGDCOWITALHYA 191
 QY 164 FODKNHLYLVMEYOPGGDLSTLNRVEDOLDENIOPYLAELIYAVSHLMGVVHEDIK 223
 DB 192 FODENHLYLVMDYVGGDLTLTSKEDKLPEDMARFYIGEMVLATDSIHQLHYHRDID 251
 QY 224 PENILVDRTGHILKLVDFGSAKONSNNKVNAKLPITGPDYMAPFVLTVANGDGKTYGLD 283
 DB 252 PDVILVDVNGHTRIDAFSGCLKNNDGTVQSSVAVGTVDYISPIILQAME-DGNGKKGPE 310
 QY 284 CDWWSGVIAIYEMITGSPFAEGTSATFNNINMFOFLKPPDD-PRVSSDFLDLIQSL 342
 DB 311 CDWWSLGVCMYEMITGSPFAESLIVETGKIMNHEERFQPSHVTQVSEAKDLIQRLL 370
 QY 343 CGGKERLKEPGL---CCHPEFSKIDMNNINSPPEFVPLTKSDDDTSNPPEPE---KNSM 396
 DB 371 CSHERRIGONGIEDPFKGAFFEGGLMNTIRNLEAPYIPDVSSPSDITSPVDVLTNTE 430
 QY 397 VSSSPCOLSP---SGFSGEELPFVGFYSKALGILGRSESVSGDPSPATSSMEKLLI 453
 DB 431 I-----LPGSHHTGFGGLHPITGTF-----TESCFSS-DGSLKLSIMQSTLT 474
 QY 454 KSEVLODSQDKCHKMEQEMTRLHRVSEVAVLSQKEVELKASGTORSLLQDLATYITE 513
 DB 475 KDEDVQ---RDLEHSLQWE---AYERRI----- 496
 QY 514 CSSLKLSEQARMEVSGEDDKALQDLHDIRQSRKLOEI-KEOGYOQVEMELMNLQLE 572
 DB 497 ---RRLEQEKLELSRKLOESTQVOSLHGSSPALSNSNRDKIKLNEIERLKKKIA 551
 QY 573 EDLVASARRSDLYESLEAESRLAAEFKRAKATECOHLLAKADQKPEVGEVANKLEKINA 632
 DB 552 D-----SNRLERQL-EDTVALLRO-----EREDS 573
 QY 633 EQQLKIQEIQLEKLEKAVKASTATELLQNIROAKERAERLEKLQNNREDSSEGRKKLVV 692
 DB 574 TQRLRGLEKQHRV-----VROKEER-----LHQGLVE 600
 QY 693 AERRRISLEKVKRLLETMERENRKLDDIOTKSOQIQOAMDKLLELEEKREAVQVSAQHL 752
 DB 601 ASER-----LKSQAKELKDAHQOKRLALQOEFSS---ELMERNAEELRAQOKV 643
 QY 753 EYVLKQKEQHYEKIKVLDNQIKKDLADKLTLENNMORHEEAHEKGIISQOKAMINAM 812
 DB 644 SRQLRKEEMEVATQKVD-----AKROEKRAE-----KLKKELEA 680
 QY 813 DSKIRLEORIIVLSSEANKLAANSLETQNNMKAQOEIMISELROQKYLETQAGKLEAQN 872
 DB 681 -----QUDDAVABASRKERKLREHSENF-----KQMESELEALKV---KQSGRGAGAT 725
 QY 873 RKLBEQLEKISHQDHDKRLLELETRLEVSLHEHEQKLELKRQLTLOLSQERSSQL 932
 DB 726 LEHQOEISIKS-----ELEKKV---LFTYEE---ELVREASHVLEVKVKEV 768

QY 933 TALQAPAALESQLOAKTELETTAAEEIQALTARDEIQKFPALRNSCTVITDLE 992
 DB 769 HDSESHQALQKETIMLKDLEKSKRRRHNMEEAVG-----TIKDKYE 812
 QY 993 EQNLQLTEDNAELNNQNFYLSKQDEASGANDEIVQARSEVDHLRREITEREMOLTSQK 1052
 DB 813 RERAMLFDENKKJLTAENEKLCSPVD-----KLTJQNR 844
 QY 1053 TMEALKTTCMTLBEQVNDLBALNDELLEKEROMEAWRSVLGDEKQFECRVERIQRMIDT 1112
 DB 845 -----QLEBLEDLDLAAK-----ESVAHNEA-----QIAELTIQWSD 876
 QY 1113 EKOSRARADQRTESRQVVELAVENKAEIITAOALKEOKLKAESLSDKLNLEKHAM 1172
 DB 877 EKD---ANGYLOALASKVTEEL-----EALRSSLSIGRTL-DPLMKV-RSQK 919
 QY 1173 LEKMAR-SLOQKLETEBELKQRLLEQAKIQO-QMDIQKHIFRLTQGLQALRADLILK 1230
 DB 920 LDMASARLELOSALEAETIRAKQLVOEELRKVYDANULTLESK-----LK 961
 QY 1231 TERSDLETOLENTIOVLVSHKVKVMEGTISQOT---KLIDPLQAKMDQPAKKKVPLOYNEL 1288
 DB 962 DSEAKRLELEMEIL---KKKMEKFRADTGKLPDPODSIPE-----YENYA 1007
 QY 1289 KLALKEKAKCAELEBEALQKTRIELRSAREEAHRKATDPHPS-----TPATARQOIAM 1343
 DB 1008 PLAMD-----LFTRTSSASEGTQAPKPEASPSWVAASEQEDM 1047
 QY 1344 SAIVRSEPHQPSAMSLAPPSRRKSSSTPEEFSRLKEMHNNIPIRFVGLMNRATKC 1403
 DB 1048 A---RPPORPSAVPL---PTQALALAGKPRX-----HQFISKSPSSPTQC 1068
 QY 1404 AVCLD-TVHGRQASKCLEQVWCHPKCSCTLPATGSLPAEYATHFLEARCRDMGNSPGL 1462
 DB 1089 SHCTSLMGLIRQYALCEVCSFACHVSCQAGAPVOCPIPE----- 1129
 QY 1463 QTKEP-----SSSLHLEGMKVPRNNKRGQOQDMRYIVLESGKYLIYD-NEAREAGQ 1514
 DB 1130 QSRKPLGVQVRGIGTAYKGVAVPKPT-GVKKGMQAVAVVCDCKLFLVLDPEGSKTOP 1186
 QY 1515 RYVEEFELCLPDDGDSIHGAVGASELANYAKADVPIYLKNEH---PHTTCWGRITLYL 1570
 DB 1189 GVIASQVLTDRDDEFVY-SSVLASDVYHATRBIDICIFRYTASILGAPSKT---SSLLI 1243
 QY 1571 LAPSFDKQWNTALE---SVVAGRYSREKABADAKLNSLLKLEGGDRDLDMNCTLPF 1627
 DB 1244 LLENENEKKKXWVIGLGLQSLIHKNRLRNQVHVPLEAYDSSLPLI---KALITTAIVD 1299
 QY 1628 SPDVVLVGTSEGLYALNVLKNSLTHVPGIAGVFOIYIKDLEKLMTAGEBALCLVDVK 1687
 DB 1300 ADRIA-VGLEBGLYIVETRDVYVRADCKKNQIELAPRKVIYILCGNNH----- 1350
 QY 1688 KVKQSLAQSHLPAOP---DISNITE---AVKCHLFGAGKIENGLCTCAAMPKSVI 1739
 DB 1351 -----HVHLVPMSSLDGAGSFIDIKLPETKGCQMLMATATILKRNSTGCTLFVAVVRLLI 1401
 QY 1740 LRYNENLSKKCIKKELETSEPCS---CIHFTNYSILIGTKNFYELDMQYTLF-----EF 1791
 DB 1402 LCYEIORTKPEFHKKNEIYVAGSVQCLAVLRDLQVQ---YSGFCLLSIQDGOPLNL 1457
 QY 1792 LDKNDLSLPAVPAASNSFPVSIYOVNSAGQREBYLLICFHEFGVFDYSYGRSRSTDILK 1851
 DB 1458 VNPNDLSLA---PLSQOSPALCAVELBS---EYILCSHNGGLYVDPGRRARARQBELM 1510
 QY 1852 WSRPLPLAFA-----YREPYLVTHFNSLEVIE-----IQARSSAGT-----PAR 1890
 DB 1511 WPAAPVACSCSPHTVTVYSBYGVDPVDRTEMWQITGLRRIRPLNSEGTILNLCPPR 1570
 QY 1891 -----AYLDINP-----RYL-----GPAISSG 1908
 DB 1571 LIYFKSKFGAGVAVNPDTSDNSKKQMLRTSKRRFVFKVPEERLQORREMLRDLPELRSK 1630

QY 1909 AYLASSYQ-----DKLRVYCKGNLYKESGTEHHRGPSTSRSSPNKRGPTTNE-- 1958
DB 1631 MISNPTFNHVAWMPGDGMVLDLPLSAVPPSQEERPGCAPPTNLA---RQPSRAKPY 1687
QY 1959 -----HITKRVASAPAPPGGSHPRRSTPHR 1985
DB 1688 ISWPGSGSEBPSVTVPILRSMSPDPQDFDKEDPSDSTYH--STPSNSSNPGGPSPNSPHR 1745

RESULT 15
ID Q9Y552 PRELIMINARY; PRT: 1711 AA.
Q9Y552
AC Q9Y552;
DT 01-NOV-1999 (TREMblrel. 12, Created)
DT 01-NOV-1999 (TREMblrel. 12, Last sequence update)
DT 01-MAR-2004 (TREMblrel. 26, Last annotation update)
DE CDC42-binding protein kinase beta.
GN Name=CD42BBP;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=99216425; PubMed=10198171; DOI=10.1006/geno.1999.5769;
RA Moncrieff C.L., Bailey M.B., Morrison N., Johnson K.J.;
RT "Cloning and chromosomal localization of human Cdc42-binding protein
kinase beta.";
RL Genomic 57:297-300(1999).
CC -1- SIMILARITY: Belongs to the Ser/Thr protein kinase family.
DR EMBL: AF128625; AAD37506.1; -.
DR HSSP: P31751; 1MKY.
DR GO: GO:0005856; C:Cytoskeleton; TAS.
DR GO: GO:0004672; F:protein kinase activity; TAS.
DR GO: GO:0007163; P:establishment and/or maintenance of cell po. .; TAS.
DR GO: GO:0007165; P:signal transduction; TAS.
DR InterPro: IPR001180; Cltron.
DR InterPro: IPR002219; DAG_Pe-bind.
DR InterPro: IPR011009; Kinase like.
DR InterPro: IPR000095; Pakbox/rhoindng.
DR InterPro: IPR001849; PH.
DR InterPro: IPR000961; Pkinase_C.
DR InterPro: IPR000719; Prot_kinase.
DR InterPro: IPR002290; Ser_thr_pkinase.
DR InterPro: IPR008271; Ser_thr_pkin_AS.
DR Pfam: PF00130; Cl_1; 1.
DR Pfam: PF00780; CNH_1; 1.
DR Pfam: PF00786; PBD; 1.
DR Pfam: PF00069; Pkinase; 1.
DR Pfam: PF00433; Pkinase_C; 1.
DR PRINTS: PR00008; DAGPDOMAIN.
DR ProDom: PD000001; Prot_kinase; 1.
DR SMART: SM00109; Cl_1; 1.
DR SMART: SM00036; CNH; 1.
DR SMART: SM00285; PBD; 1.
DR SMART: SM00220; S_TKc; 1.
DR SMART: SM00133; S_TK_X; 1.
DR PROSITE: PS00106; CRIB; 1.
DR PROSITE: PS00479; DAG_Pe_BIND_DOM_1; 1.
DR PROSITE: PS50081; DAG_Pe_BIND_DOM_2; 1.
DR PROSITE: PS50003; PH_DOMAIN; 1.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE: PS00106; PROTEIN_KINASE_ST; 1.
DR ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
KW SSOURCE 1711 AA; 194356 MW; 3A3731D40A363497 CRC64;

Query Match 13.7%; Score 1438; DB 2; Length 1711;
Best Local Similarity 24.4%; Pred. No. 1e-33; Mismatches 668; Indels 584; Gaps 76;
Matches 512; Conservative 336;

QY 44 SPLSRGCIIDALVLTREESQPLMKIKVNSNVRKXSTIATLOELQPSAKDPEVRSIV 103
DB 23 SALSVELTLDLVLCITYECSHSAIRDRDYAELENAKPTQIVKMQIREDDEFIIVKI 82
QY 104 GCGHPAEVQVREKAGDIIYAMKVMKKKALLAEOVSFPFEENIISRSTSPMIPOLQYA 163
DB 83 GRAPGEVAVAKKNTERTIYAMKILNKMKLAETACFEERDVLVNGCCQITALTHYA 142
QY 164 FQDKNHLVYVMEYQPGDILLNRYEDQDENIIOFYLAELLIVASVHLMGVNHRDIK 223
DB 143 FQDENHLVYVMDVYVGGDILLTLKFKEDKLPEDMARFYIGEVLAIDISIQHLYVHRDIK 202
QY 224 PENILVDRTGHILVDFGSAKNSKMNNAKPIGTPTDMAEVLTVWNGDKGTGGLD 283
DB 203 PDVLLDVNHGHIILADFGSLKNNDGTQVSSVAVGHPDIYSBELLQAMM-DEMGKKGP 261
QY 284 CDMAVGVLAEMITVGRSPAEGTARSATENNINNFORFLFPPD-PRVSSDFLDLIQSL 342
DB 262 CDMAVGVCMYEMITVGRSPAEGLVETTYGKINMHERFQPSHVTVDVSEADLIQRLI 321
QY 343 CGQKERLKEFGL---CCHPFSKIDNNINRNSPPVPVTLKSDDTSNDEPE---KNSW 396
DB 322 CSRRERLQNGIDPFKHAFFEGGLNMENIRNLEAPYIPDVSSPSDTSNFDVDDVLRNTE 381
QY 397 VSSSPCOLSP---SGRSGEELPVGFSYSKALGILGRSESVSGLDSPATSSMEKKLTI 453
DB 382 I-----LPGSHITGSGHLPIGTFT-----TESCFS--DRGSLKSTMQSNTLT 425
QY 454 KSKELDSDQCKHMOEMRILHRVSEVAVLSQKEVELKASGTQRLLEQDLATYITE 513
DB 426 KSDVQV--RDLEISLQNE--AYERRI----- 447
QY 514 CSSLKSRLSEQARMEVSEQEDKALQLHDIRQSKLOEI-KEQGYQAYQVEEMLMNQL 572
DB 448 -----RLREQKELERKIQESTQVQSLGSSRALSNSRDEIKYKINAEIRLNKXIA 502
QY 573 EDIVSARRSDLYESLRSRLAAEFKRAKTCQHKLKAKQOGKPEVGYAKLEKINA 632
DB 503 D-----SNRLERQI-EDTVLRQ-----EREDS 524
QY 633 EQLQKIQEOLQEKLEKAVKASTATELLQNTROKERAERLEKTLQNRSDSSEGRKKLVE 692
DB 525 TORLRLGJEKHVR-----VQKEKE-----LHKQV 551
QY 693 AERRRSLNKVYKRLTMRERENRLKDIQTSQOIQOMADKILBEKREAOVSQHL 752
DB 552 ASER-----LKSQAKELMDAHOQRKALQERS-----ELNERVAELRAQOKV 594
QY 753 EVHLKQEQHYEEKIKVLDNQIKDLADKETLENMQRHEEBAHEGKILSEKAMINAM 812
DB 595 SRQLRDKKEEMEVATQVD-----AMRQEMRRAE-----KLRKEDEA----- 631
QY 813 DSKIRSLQRIIVLESAKTLAANSILFTQRMMAQOEMISLQOKPYLETQAGKLEAON 872
DB 632 -----QDDVAVAASRKRLRHSNFC-----KQMSSELALKV---KQGRGAGAT 676
QY 873 RLKEQLLEKISHODSDKNLLELTRLEVSLEHSEOKLELRQUTLEQLSQEBSQ 932
DB 677 LEHQEISIKS-----ELEKKV---LFYEE---ELVRREASHVLBYKAVKEV 719
QY 933 TALQAPAALESQRLQAKTELEETVABEEIQTALTAHRDEIQRKDALRNSCTVITDLE 992
DB 720 HDSESHQALQKEIILMLKDLKESKERNHEMEAVG-----TIQKYE 763
QY 993 EQNLQITENAEILNNQNFYLSKQLDASGANDBIYQLRSVHDHARREITEREMQLTSQK 1052
DB 764 REBAMLPDEKKKTLAENEKLCSTVD-----KLTQKOR 795
QY 1053 TMEALKTCTMEEQVMDLEALNDELLEKERQWEAMRSYLGDEKQFECRVRELQRLDT 1112
DB 796 -----QLEDELQDLAAKK-----ESVAHWEA-----QIAILIQWVSD 827
QY 1113 EKOSRAPADQRTESQVVELAVKEHKAETLALQALKEQKLKAESLSDKLNDLEKKHAM 1172

Db 828 EKD--ARGYLOALASKYTEEL-----EALRSSSIGSKRTL--DPLMKV-RRSQK 870
QY 1173 LEMNAR-SLOOKLETERELKORLLEBQAKLOO-CMDLOKNHIFRLTQIGLOALDRADLLK 1230
Db 871 LDMGARLELOALBEMELRAKQLOVQELKVKDANLLESK-----LK 912
QY 1231 TERSDLEYOLENIQVLSHEKVKMEGTISQOT--KLIDFLQAKMDQPAKKKKVPLQYNEL 1288
Db 913 DSBANKRELLBEMEL-----KKMEKFRADTGLKLPDQDSIFE-----YFNTA 958
QY 1289 KLALEKAKACALEBEALQKTRIELRSAREBAHRKATDHPHS-----TPATAQOIAM 1343
Db 959 PLADH-----LTERTSASEQETQAPKPEASPEMSVAASEQEDM 998
QY 1344 SAIVRSPHOPSAMSLAPPSRRKESSTPEFSRRLKERMHNIPIHRFNVGLNMBATKC 1403
Db 999 A-----RPPQBSAVPL--PTTQALVLAGPKKA-----HQSISKSFSSTQC 1039
QY 1404 AVCLD-TVHFGROASKLEQCVMCPCSTCLPATGCLPAEYATHTFAFCRDKXNSPGL 1462
Db 1040 SHCTSLMWGLIRQYACGVCSFACHVSCDKGAPQVCPIRPE----- 1080
QY 1463 QTKEP-----SSSLHEGMMKVPRNNKRGQCGWDRKIYVLEBSKVLID-NEAREAGQ 1514
Db 1081 QSKRPLGVDVORGIGTAKYGHVKPKPT-GYKKGWQRAYAVWCECKFLYDLPEGSTOP 1139
QY 1515 RPVEFELCLPDGVSHGAVGASELANATAKADVYLKMSH-----PHTCWPGRTYL 1570
Db 1140 GVIASQVLDLDEFSV--SSVLASDVHATRDIPCIPTVATSLGAPSKT---SSLLI 1194
QY 1571 LAPSPDKORWVTALE---SVVAGRVSRERAEADAKLGNLSLKLBSGDRLDMMCTLPF 1627
Db 1195 LTNENEXKXKXVGLIEGLQSLHKRLNQVHVPLEAVDSSLPLI---KAILTAIVD 1250
QY 1628 SDQVVLVTEBGLYALANVKNSLTHVPGIGAVFOYIYIKLEKLIMAGEERALCLVDK 1687
Db 1251 ADRIA-VGLEBGLYIVETRIADVIVAAADCKVHQIETLAPREKIYVILCGRNH----- 1301
QY 1688 KVKQSLAQSHLPAP---DISPNIFE---AVKGCHLFGAGKIENGLCICAMPKXVI 1739
Db 1302 -----HVHLYPMSGLDABEGSFDIKLPETKGCQMLATATLKRNSGTCLFVAVKRLI 1352
QY 1740 LRYNENLSKYCIRKEIETSEPCS--CIHFTVYSILIGTNKFEIDMKQYTL-----EF 1791
Db 1353 LCYEIQRTKPRHRKENEIVAGSVQCLAVLDRLCVG---YPSGFCLLSIOGDOPLNL 1408
QY 1792 LDKNDHSLAPAVFAASSNSFPVSIQVNSAGQREBYLLCFHEFGVFDVSYGRSRRTDJK 1851
Db 1409 VNPNDPSLA---FLSQGFALCAVELS---BEYLLCFSHMGLYVDPOGRARAQELM 1461
QY 1852 WSRLLPLAFA-----YREPYLFTYHNSLEVIE-----IQARSSAGT-----PAR 1890
Db 1462 WPAAPVACSCPTHTVTVSEYGVDFVRTWEMVOTIGLRIRPLNSGTINLNLNCEPFR 1521
QY 1891 -----AYLDIPNP-----RYL-----GPAISSG 1908
Db 1522 LIYFKSKRSGAVLANVPTSDNSKCOMLSTRKRRFVFPVEBERLOQRREMLRDBELRSK 1581
QY 1909 AIYLASSYO-----DKLRYICCKGNLVKESGTEHHRGPTSRSSPNKRGPPTYNE-- 1958
Db 1582 MISNPTNHNVAHMGPGQMQLMDLPLSAVPPSQEERPGAPNTLA---RQPERNKPY 1638
QY 1959 -----HITKVASPAPPEGSHPRESTPHR 1985
Db 1639 ISWPSGSGSEPSVTVPLRSMSPDQDFDKPPDSSTKH--STPSNNSNPSPGPPSPNSPHR 1696

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 11, 2005, 16:32:13 ; Search time 63 Seconds
(Without alignments)
3136.972 Million cell updates/sec

Title: US-10-791-666-2

Sequence: 1 MLKFKXGANNPLDAGAAEPT.....AGAVRTPLSQVKNVMDQSSV 2054

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	7935	75.6	1597	2	citron - mouse
2	1524.5	14.5	1732	2	protein kinase (EC
3	1419.5	13.5	1702	2	protein kinase (EC
4	1283.5	12.2	1548	2	hypothetical prote
5	1275.5	12.2	1354	2	serine/threonine-s
6	1263.5	12.0	1388	2	serine/threonine-s
7	1260	12.0	1354	2	serine/threonine-s
8	1250.5	11.9	1388	2	serine/threonine-s
9	1126.5	10.7	1173	2	hypothetical prote
10	888.5	8.5	524	2	protein kinase (EC
11	872	8.3	557	2	serine/threonine-s
12	761	7.3	1356	2	hypothetical prote
13	751	7.2	1286	2	hypothetical prote
14	728	6.9	1256	2	hypothetical prote
15	721	6.9	522	2	protein kinase (EC
16	715.5	6.8	479	2	protein kinase (EC
17	710	6.8	441	2	myotonic dystrophy
18	710	6.8	474	2	myotonic dystrophy
19	710	6.8	516	2	myotonic dystrophy
20	707	6.7	756	2	probable protein k
21	701	6.7	526	2	protein kinase PKT
22	700	6.7	469	2	serine/threonine-s
23	697	6.6	596	2	probable protein k
24	696	6.6	569	2	hypothetical prote
25	676.5	6.4	443	2	probable protein k
26	676.5	6.4	475	2	protein kinase [m
27	674.5	6.4	465	2	protein kinase (EC
28	670	6.3	500	2	protein kinase (EC
29	660.5	6.3	483	2	protein kinase P41

30	639	6.1	545	2	T01288	protein kinase F27
31	636.5	6.1	665	2	S70706	probable protein k
32	636	6.1	480	2	T47255	serine/threonine k
33	636	6.1	598	2	T47254	serine/threonine k
34	636	6.1	620	2	S22711	probable protein k
35	631.5	6.0	412	2	T78395	myotonic dystrophy
36	608	5.8	908	2	T25035	hypothetical prote
37	604	5.8	1099	2	A56155	tumor suppressor p
38	591.5	5.6	1909	2	A45592	liver stage antiage
39	580	5.5	1410	1	A57013	early endosome ant
40	574	5.5	2139	2	T18296	myosin heavy chain
41	573.5	5.5	1738	2	T14667	interactin - slime
42	567.5	5.4	1790	2	S67593	transport protein
43	550.5	5.2	2442	2	T08621	centrosome associa
44	549.5	5.2	624	2	T41341	probable serine-th
45	548	5.2	607	2	S62556	probable serine/th

ALIGNMENTS

RESULT 1
S68420
citron - mouse
C/Species: Mus musculus (house mouse)
C/Date: 15-Feb-1997 #sequence_revision 13-Mar-1997 #text_change 09-Jul-2004
C/Accession: S68420
R/Madaule, P.; Furuyashiki, T.; Reid, T.; Ishikaki, T.; Watanabe, G.; Morii, N.; Narumi,
FEBs Lett. 377, 243-248, 1995
A/Title: A novel partner for the GTP-bound forms of rho and rac.
A/Reference number: S68420; MUID:96128238; PMID:8543060
A/Accession: S68420
A/Status: preliminary; nucleic acid sequence not shown
A/Molecule type: mRNA
A/Residues: 1-1597 <MAD>
A/Cross-references: UNIPROT:P49025; EMBL:U03904; NID:G1079733; PIDN:AACS2341.1; PID:G107
C/Keywords: alternative splicing
F/911-979/Domain: protein kinase C zinc-binding repeat homology <K22>

QY	468	MEQEMTRLHRRVSEVAVLISQKEVELKASETORSLLEQDIATYITTECSSLKRSLEQARME	527
DB	9	MEQEMTRLHRRVSEVAVLISQKEVELKASETORSLLEQDIATYITTECSSLKRSLEQARME	68
QY	528	VSQEDDKALQLHDIREQSRKLEIYKQEQVQAVEEMRLMNNQLEEDLVARRRSDLYES	587
DB	69	VSQEDDKALQLHDIREQSRKLEIYKQEQVQAVEEMRLMNNQLEEDLVARRRSDLYES	128
QY	588	ELRESRLAEEPRKATKTECHKLLKAKDQKPVGAYKLEKTNAAQQLKIOLEQLEK	647
DB	129	ELRESRLAEEPRKATKTECHKLLKAKDQKPVGAYKLEKTNAAQQLKIOLEQLEK	188
QY	648	AVASTATEATELLQNIROAKERARELEKLNREDSEGRKLVLEAEERHSLLENKVKRL	707
DB	189	AVASTATEATELLQNIROAKERARELEKLNREDSEGRKLVLEAEERHSLLENKVKRL	248
QY	708	ETMERRENRKLDIQYKSQOIQMADKYLELEKHBDAQVSAQHLVHLKQEQHYBEKT	767
DB	249	ETMERRENRKLDIQYKSQOIQMADKYLELEKHBDAQVSAQHLVHLKQEQHYBEKT	308
QY	766	KVLNDQIKKDLADKETLEMMQHEEAEHKGKILSEQAMTNAMSKRSLEKRVLS	827
DB	309	KVLNDQIKKDLADKETLEMMQHEEAEHKGKILSEQAMTNAMSKRSLEKRVLS	368
QY	828	EANKLAANSLFTQRMKAQEEWISLRQOKFLFTQAGKLEAQNKLLEBOLKISHOH	887
DB	369	EANKLAANSLFTQRMKAQEEWISLRQOKFLFTQAGKLEAQNKLLEBOLKISHOH	428
QY	888	SDKNRIELETRILEVSLSEHEQKLELKRQLTLEQLSIOERESQLTALQAAAPALESQLR	947

```
Db 429 SDSRLLEETRIREVSLEHEBOKELKQUTELQSLQREBSQTLQAPRALBSQLR 488
Qy 948 QAKTELEETTAABEEBEOALTARHDEIQRFKDALRNSCTVITDLBEQANQJTEBDNALNN 1007
Db 489 QAKTELEETTAABEEBEOALTARHDEIQRFKDALRNSCTVITDLBEQANQJTEBDNALNN 548
Qy 1008 QNFYLSKQDLDEAGANDEIVQASVNDHREITEREMQUTSQQYQTEALKTCTMLEEQ 1067
Db 549 QNFYLSKQDLDEAGANDEIVQASVNDHREITEREMQUTSQQYQTEALKTCTMLEEQ 608
Qy 1068 VMDLEALNDELKEEROMEARNSVLGDEKSOFECEVRELQRMJLDTKOSPARADQRTTES 1127
Db 609 VMDLEALNDELKEEROMEARNSVLGDEKSOFECEVRELQRMJLDTKOSPARADQRTTES 668
Qy 1128 ROVVELAVKHEKAETIALQOALKEQKKAESLSDKLNDELKQAMLEMNARSLQOKLETE 1187
Db 669 ROVVELAVKHEKAETIALQOALKEQKKAESLSDKLNDELKQAMLEMNARSLQOKLETE 728
Qy 1188 RELKORLEBOAKLQOQMDLQKNHIFRLTQGLQALDRADLLKTERSDLEYOLENTQVLY 1247
Db 729 RELKORLEBOAKLQOQMDLQKNHIFRLTQGLQALDRADLLKTERSDLEYOLENTQVLY 788
Qy 1248 SHEKVMKEGTISQOTKIDFLQAKMDQPAKKKVPLQYNELKJLALKEKARCAELEBALQ 1307
Db 789 SHEKVMKEGTISQOTKIDFLQAKMDQPAKKKVPLQYNELKJLALKEKARCAELEBALQ 848
Qy 1308 KTRIELRSAREBAHRAKATDHPHSTPATARQOLAMSAIVRSPEHQPAMSLAPSSRR 1367
Db 849 KTRIELRSAREBAHRAKATDHPHSTPATARQOLAMSAIVRSPEHQPAMSLAPSSRR 908
Qy 1368 KESSTPEESRRLKERMHNI PHRFVNGLMARATKCAVCLDTYHFGQASKLECYQVCH 1427
Db 909 KESSTPEESRRLKERMHNI PHRFVNGLMARATKCAVCLDTYHFGQASKLECYQVCH 968
Qy 1428 PKCSTCLPATCGLPAAEYATHTFAFCQDKNSPGLQTKEPSSSLHEGMMKVPRNNRGQ 1487
Db 969 PKCSTCLPATCGLPAAEYATHTFAFCQDKNSPGLQTKEPSSSLHEGMMKVPRNNRGQ 1028
Qy 1488 QGMDRKXIVLEGSKVLVYDNEABAGQRYEPEELCLPDQVSIHGAVGASELANAKAD 1547
Db 1029 QGMDRKXIVLEGSKVLVYDNEABAGQRYEPEELCLPDQVSIHGAVGASELANAKAD 1088
Qy 1548 VPIYLLKMEHSHPTTCWPGRTLYLLABSPDKOKWYTLLESVAGAGRSRREKAEADALG 1607
Db 1089 VPIYLLKMEHSHPTTCWPGRTLYLLABSPDKOKWYTLLESVAGAGRSRREKAEADALG 1148
Qy 1608 NSLLKLEGGDRLLMNCTLPSPDOVLYVGTBGLYALVNLKNSLTHVHGIGAVFOIYIKD 1667
Db 1149 NSLLKLEGGDRLLMNCTLPSPDOVLYVGTBGLYALVNLKNSLTHVHGIGAVFOIYIKD 1208
Qy 1668 LEKILMIAGEERALCLVDYKVKQSLAQSHLPAQPDISPNIPEAVKQCHLFGAGKLENGL 1727
Db 1209 LEKILMIAGEERALCLVDYKVKQSLAQSHLPAQPDISPNIPEAVKQCHLFGAGKLENSL 1268
Qy 1728 CTAAMPKSVIILRYNENLSKYCIKEIETSEPCSHFTNYSILIGTKNFYEIDMKQYT 1787
Db 1269 CTAAMPKSVIILRYNENLSKYCIKEIETSEPCSHFTNYSILIGTKNFYEIDMKQYT 1328
Qy 1788 LEPEFLDNHSLAPAAVAASSNSPVSIVQANSAGQEEYLLCHEHGVGVDSGRRSRT 1847
Db 1329 LEPEFLDNHSLAPAAVAASSNSPVSIVQANSAGQEEYLLCHEHGVGVDSGRRSRT 1388
Qy 1848 DDLEKMSRLPAPAREPYLFEVTHNSLEVIEIQARSAGTPARAYLDIPRPRYLGAPISS 1907
Db 1389 DDLEKMSRLPAPAREPYLFEVTHNSLEVIEIQARSAGTPARAYLDIPRPRYLGAPISS 1448
Qy 1908 GAIYLLASSYQDKLVICCKGNLVKESGTEHHRGPTSSPNNKGPPTVNEHTKRVASS 1967
Db 1449 GAIYLLASSYQDKLVICCKGNLVKESGTEHHRGPTSSPNNKGPPTVNEHTKRVASS 1508
Qy 1968 PAPPEGSHRPEBTPRKY--REGRTLRBDKSGFRLERKSGRILTSRRERSPARLF 2025
Db 1509 PAPPEGSHRPEBTPRKY--REGRTLRBDKSGFRLERKSGRILTSRRERSPARLF 1568
```

```
Qy 2026 EDSSRGRPLPAGAVRTPLSQVNNKWDQSSV 2054
Db 1569 EDSSRGRPLPAGAVRTPLSQVNNKWDQSSV 1597

RESULT 2
T14039
protein kinase (BC 2.7.1.37), myotonic dystrophy-associated - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 20-Sep-1999 #sequence_rev150n 20-Sep-1999 #text_change 16-Aug-2004
C:Accession: T14039
R:Jung, T.; Chen, X.O.; Tan, I.; Manser, E.; Lim, L.
Mol. Cell. Biol. 18, 130-140, 1998
A:Title: Myotonic dystrophy kinase-related Cdc42-binding kinase acts as a Cdc42 effector
A:Reference number: 217862; PMID:98078670; PMID:9418861
A:Accession: T14039
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1732 <LEU>
A:Cross-references: UNIPROT:O54874; EMBL:AF021935; NID:g2736150; PID:g2736151; PIDN:AACO;
C:Genetics:
A:Gene: MRCK
C:Superfamily: protein kinase homology
C:Keywords: ATP; phosphotransferase
F:/5-343/Domain: protein kinase homology <KIN>
F:/1013-1062/Domain: protein kinase C zinc-binding repeat homology <KZN>

Query Match 14.5%; Score 1524.5; DB 2; Length 1732;
Best Local Similarity 24.8%; Pred. No. 6.9e-34;
Matches 529; Conservative 335; Mismatches 667; Indels 599; Gaps 76;

Qy 36 PFMTQOQMSPLSREGILDALFVLPFEECSQPALMKTHVSNFVRKYSDTTAELOQPSAK 95
Db 18 PAQTNQCC--FSVETLLDILICLYDECNNSPLREKRNIILEYEMAKPFSTKVQMBLHRE 75
Qy 96 DFEVRLVNGGHAFAVQVUREKATGDIYAMKKKKALAOEVSFPEERNILSSSTSP 155
Db 76 DFEILVYIGGAGAEVAVVKLKNADKVPFAKILINKKEMLGAEATACREERDVLVNGDSK 135
Qy 156 WIPOLYAFQDKNHLVMEYQPGDILLNRYEPDQDENLQFYLAELILAVSHVLM 215
Db 136 WITTLATYADDDNNLLVNDVYVGDLLTLTSSFEEDRLPEMARFVLAEMVAILDSVHQL 195
Qy 216 GYVHARDIKENILVDTGHIKLVDFGSAAMKSNKRVNAKLPIGTDPYAAEVLTVMGCD 275
Db 196 HYVHRDIKPDNILLMDNGHRLADFGSCLMEDGTQGSVAVGTPDYISPEILQAME-D 254
Qy 276 GKGTGLDCDQMSVGYIAYEMITGRSPFAEGTARFPNNIMNFQRLKRPDD-PKVSSDF 334
Db 255 GKGRYRPECQMSVLGYCMYEMLVGETPFAESLVETGYKIMHKERRPQPTQVTVDSNA 314
Qy 335 LDIQLSLCGQKRLKFEGL---CCHPFPSKIDMNNIRNSPPFVPLTKSDDTSNFDEP 391
Db 315 KDLIRRLIGREHRLQONGIEDPKRPFPSGIDMDIIRNCEALYIIEVSPDTTNSFVDV 374
Qy 392 E---KNSWSSSPQQLSPGSGEELLPFVGFYSKALGILGNS-----ESV 434
Db 375 DDLCKNSEMTPEP---THAFSGHHLPFVGFTYTSQVLSDRCLRTAGPTSLDIDVNV 431
Qy 435 VSGLDSPAKTSSMEKLLIKSKELQDSQDKCHMKQEMQRLHARVE---VEAV-LSQK 489
Db 432 QRTLDNNLTALEARIRI-----KRLDEKELTKLDEESTQTVQALQYSTV 477
Qy 490 EVELKASETQSRSLBEDLATYITEGSSLKRSLEQAMEVSEQEDKALQLLHDIREQSRKL 549
Db 478 DGPLTAS-----KDL-----EIKSLKEIIRKAKQVAENV-----507
Qy 550 QEIKEOEYQAYVEEMLMNQLEEDLVASARRSDLYESILRESRLAAEFKRYATECQHK 609
Db 508 -----HLEQQLBE-----ANSVARELDADAFRQIKAREKQIKTLQOREEELINKELVQASER 557
Qy 610 LLAQKQGRPEVGEYAKLEKINAEOQLKIQELOEKLEKAVKASTAETELLQNIROAKERA 669
```

```

Db      558 L---KNSK-----ELKQHCQKLAHQFMEINERTELTHQKQALAHVNDKEEV 607
Qy      670 ERELEKLONRDSESGIRKKLVAEERHSLNKKLETREERNLKODIQKVSQOIQ 729
Db      608 DLVWQK-----AESIRQELRAERAKKEVHTHEALIAASQKRLREGRHSYKQLE 660
Qy      730 QMADKILEBKEHREAOVSQAHLEVLKQKQHYBEKIKVLDNQ-----IKKDLADKETT 784
Db      661 N-----ELF-----GLKQKQISYSPGCSIEHQOETIKKLTDEKKSIF 699
Qy      785 ENMMQREBEAHEKGLISEQKAMINMDSKIRSELQRIVELSEANKLAANSLSFTQRM 844
Db      700 -----YEEISKKEGJHA-----SEIKNLKKELDH-SGQOQALAKKEIMV----- 738
Qy      845 KAOEWMSIELROQKYLETQAGKLEAQRKLEOELEKISHODSHDKNRLLEFTRLEVS 904
Db      739 -----LKDLEK-----TR-RESQ 751
Qy      905 LEHEQKLEKQKQTLQELQSLQERESQALQAAAPALLESQLRQAKTELETTAEAEEI 964
Db      752 SEREEPENEFEQO-----YEREKVL----- 771
Qy      965 QALTANDEIQKRDALRNSCTVITDLEBQNLQTEDNAELNQNPFYLSQDLASGAND 1024
Db      772 -----LTERNKKLT----- 780
Qy      1025 EIVOLRSEVDHLREITEREMOLTSQKQTEALKTCTQMLEEQVMDLEALNDELLEKRO 1084
Db      781 -----SELD-----KLTS-----LYESLSLRQHLEEKVOLA-----DKKESVAH 816
Qy      1085 WEAMRSVLDGSKQFECRVELQRMIDTEKQSPARADQRTESQVVELAKVHKAEILA 1144
Db      817 WEA-----QITEIILQWVSDEKQAR-----YLQALA 842
Qy      1145 LQQAALKEQKAKBSISDKLNDLE---KKHAMLENAR-SIQQKLETERELKORLLEQAK 1200
Db      843 SKRTEBELAERNSLSGTRATDMKMRKRFALDLMSALEQSALEDAIRAKQALQOEILNK 902
Qy      1201 LQOQMDLQKNHIFRITQLOALDRADLKTERRSDLEYOLENIOVLASHEKMEGTISQ 1260
Db      903 VK-----ASNITTECKLXQSEKKNLLESEIBQILIDTELEBKEKVEH 946
Qy      1261 QTKLIDFLQAKMDQPAKKKKVPLQYNELKLALEKAKCAELBEALOKTRIEBLSAREEA 1320
Db      947 RDSQHSPL-AFLNTPF-----DALDQFE---RS----- 970
Qy      1321 AHRKATDHPHS--TPATARQOIAMSAIVRSBEHOPSAIAPSSRRKSSSTPEEFSRR 1379
Db      971 -----PSCPPACKGRRIADSAFL--PVHTPFL-----RKKGCPASAGFPK 1009
Qy      1380 LKERMHNIHFRFVGLNMRATKCAVCLD--TVHGRQASKCLEQWQCHPKCSTCLPATC 1438
Db      1010 RK-----THQFFVVSFTAPTKCHOCTSLMVGILRQGCSEVGSFCHTTCVNAKAPTC 1062
Qy      1439 GLPAEYATHTFEAFCDKXMSPGIQTPE-----SSSLHLEGMKVPRNNKRG--COGW 1490
Db      1063 PVPEPE-----QTKGPIGIDQKQVGTAYBCHVAILPK--PAGVAKGW 1101
Qy      1491 DRKTYLESGKVLTYDNEAREAQO--RPVEBELCLPGDVSIHGAVGASIELANTAKADV 1549
Db      1102 QRALAVVCPKFLFYLDIAEGKASQSSVISOVIDMRDEEFSV--SSVLASDVIIHARSIDIP 1160
Qy      1550 YILKME-----SHPHTCWPGRTLYLLAPSPDKQWMTALESVAVAGRVREKAEADAKL 1605
Db      1161 CTRFVAVASQLSAPSDK-----STIMLADSETERSKMWVISEL--HKVLAKKXKFRRSV 1213
Qy      1606 LGSNLSLKEGDDRLDMNCTLP-----PSDQVVLVGTREGLYALNKLSTLTPPG 1655
Db      1214 -----YVPKEAYD-----STLPLIKTTQAAAIIOHERVALGNESGLFVHTHKOELLIRVD 1264
Qy      1656 IGAIVQYIITKDLBKLMINGEBRALCLVYKVKQSLAOSHILPAOPDISPNIPEAVKGC 1715

```

```

Db      1265 NKTIHQELIPSDQVAVISGRNRHVLFFPMSALDGETDYYKAE-----TKGC 1314
Qy      1216 HLFAGAKIENGL--CICAMPKSVVILRYNENLSKYCIKEIETSPSCIHFTNYS--I 1771
Db      1315 QTIAGKVRHIALSLCLVAMKRGVLCYELFQSKTRHKKFEIQV--PCNVQMAIFSEHL 1372
Qy      1772 LIGTKVEYELDMKYTLE-----EFLDKXDSLAPVAFSAASSFPVSIYQVNSAGRE 1825
Db      1373 CVG-----FQSGFLRYPPLNGEGSPCNMLHSDHTLA---FITHQPMDAICAVEISN---K 1421
Qy      1826 EYLICFHEFGVNDVSYGRRSRTDCLKWSRLPLAVAREPYLFVHFNLSLEYEIQARSSA 1885
Db      1422 EYLICFSSIGIYDCCGRSRQOELMPANPSSCCYAPLSTISENAVDIFVNSMEWI 1481
Qy      1886 GTPARAVLIDIPNPRYLCPAISGAIYVASSYODCLARYICCKGNLVKES-----GTEHRRG 1940
Db      1482 QF-----LPLKVR---PLMTGSSINLG--LETRILYFNKRAEDELVVPETSDNR 1531
Qy      1941 PSTSRSSPNKR-----GPPTYNE-----HI 1960
Db      1532 KQWRNINNNKRRYSFVPEEERMQRREMLDPEMRNKLISNPFNFHIAHMGDGIQI 1591
Qy      1961 TKGVASSPAPPEG-----PSHREPESTPHRYREGRTEL--RDKSRRGLEREKS 2008
Db      1592 LKDLPMNPRQESRTVFSVSIPSTIKSRPEPGRSMSSGSLARSQAONGALKRHS 1651
Qy      2009 PGRILSTRERSPARLFEDESSRGLPAGAV 2038
Db      1652 GGS-YNTKQPMBS-----PSEGLSSGCV 1675

RESULT 3
T14050
protein kinase (BC 2.7.1.37) beta, myotonic dystrophy-associated - rat
C:Species: Rattus norvegicus (Norway rat)
C>Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 16-Aug-2004
C/Accession: T14050
R/Leung, T.; Chen, X.Q.; Tan, I.; Manser, E.; Lim, L.
Mol. Cell. Biol. 18, 130-140, 1998
A:Title: Myotonic dystrophy kinase-related Cdc42-binding kinase acts as a Cdc42 effector
A:Reference number: Z17862; MIMD:98078670; PMID:9418861
A/Accession: T14050
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A/Residues: 1-1702 <LEU>
A:Cross-references: UNIPROT:O54875; EMBL:AF021936; NID:g2736152; PID:g2736153; PIDN:AACO
A:Experimental source: brain
C:Genetics:
A:Gene: MRCK-beta
C:Superfamily: protein kinase homology
C:Keywords: ATP; phosphotransferase
F:7-342/Domain: protein kinase homology <KIN>
F:1027-1076/Domain: protein kinase C zinc-binding repeat homology <KZN>

Query Match 13.5%; Score 1419.5; DB 2; Length 1702;
Best Local Similarity 24.7%; Pred. No. 4.2e-31;
Matches 482; Conservative 310; Mismatches 623; Indels 533; Gaps 64;

Qy      44 SPLSRGILDALVYLEEGSCQPALMKIVSNFYRKYSDTITAELOEQPSAKDFVRSIV 103
Db      23 SLSLSEVTLIDVLCLYTECHSALRRDKYVAEFLMAKPFTOVKMDQLHREDEIILKVI 82
Qy      104 GCGHPAEVOVVRKATGCDIYAMKVMKKKALLAEOVSFFEEERNIISRSNPTPOLQYA 163
Db      83 GRGAPGEVAVAKKNTERTIYAMKILNKWEMLKAEFTACFEERDVLVNGCQWITLALHYA 142
Qy      164 FQDKNHLVYVMEQPGDDLSTLRRYEDQDENLQFYELNELILAVSVHLMGVHNDIK 223
Db      143 FQDENLYLVMDYVYGDDLTLTSLKPEDKLPEDMARRYIDEMVLADISIHQLNHHVDIK 202
Qy      224 PENILYDRGTGHIKLVDFGSAAKNSKNQVNAKPIGTGPDYMADEVLTVMNGDGKGTYGLD 283
Db      203 PDNVLLDVNGHIRLADFGSLKNNDDGTQSSVAVGTPDYISPEIIQAME-DGWKGYGPE 261

```



```

Db      19 NITWDG-----PSKKPRLSPFTLIDSLICLYDECNSYLRKEKCIAPFESVKTIVLSKA 73
Qy      88 QELPQSAKDEEVSRLVCGHFAEVOVREKATGDIYAMKWKKKALLAOEQVFFEEERN 147
Db      74 KKLRLSDDEPVLKVIKGAFAEVAVVRMGVGEIYAMKILNKEMVRAETACFRREERD 133
Qy      148 ILSRSTSPWITPOLQYAFQDKNHLVLYMEYOGGDLISLNRVEDQDLENTIQPYLAELIL 207
Db      134 VLVYGDRRMTITNLHYAFQDEKNLTFYMDYYIGGDMTLISKFVHIPESMKFIAMVIL 193
Qy      208 AVHSVHMGVYVHDIKENILVPTGHIKLVDPSSAKMNSKNVNAKLPITGTDVNAPE 267
Db      194 AIDSLHRLGYVHRVDPKDNVLLDMQGHIRLADFGSCIRILADGSAVNAVAGTDDYISPE 253
Qy      268 VLVYMGNDGKGTGYLDCDMMSVGIAYEMITGRSPFAEGTSARTFNNIMNFQRLKPPDD 327
Db      254 ILKAME-DGRGRYCKECDMWSLIGICMYEMLYGTTTFPSERLVDTYGIKIMQMDLPDD 312
Qy      328 P---KVSSDFLDLIQSLCGQKERLKEGLC---CHPFPKIDMNNIRNSPPFPVTLKS 381
Db      313 EIDMVVSEAKDLIRQLICSSDVAFGHNGLSDFQLHFPFGIDMNTIRDSNPVPVPEVS 372
Qy      382 DDDTSNDEPERKSWSSPCQLSP-----SGFSEELPVPVGFYSKALGILGRESVVG 437
Db      373 PEDTSNFDVVCED--DFTPCETQPPRVLAFTGNHLPPVGFSTYHG--SLISDARSILDE 429
Qy      438 LDBPAKTSMEKKLLIKSKELQSDQDKHMOEMLTHRRVSEVAEVLISKEVELKASE 497
Db      430 IRLAIQRCQDAELMEKSV-----GPMLELENKAEIVQKAKAQITTIKQHVANRSE 484
Qy      498 TORSLEQDLATYITTECSSLKRSLSEARMEVSQEDKALQILHDIRQSRKLOEIKOEY 557
Db      485 EDEN-----YESTIAQK-----497
Qy      558 QAQVEEKLMMNOLEEDLVARRRSDLYESELRESRLAEEFRKATECQHKLIKADQG 617
Db      498 ----DEIQIINKLEDEAL-----AQOQO 517
Qy      618 KPEVGEYAKLEINAEQOLKIOLQELQEKAYKASTATELLQIRQAKRABERELKQ 677
Db      518 KPR-----DELYASEKKLEKELKERNKOLV-----MEKSEIQLELDNIN 556
Qy      678 NREDSSEGIKKVLEAEERHSHLENKYRLLETMERERNRLKDIQTSQOIQWADKILE 737
Db      557 DHD-----QVLYEKATVVQOR-----DMQA---ELADGDSLL- 588
Qy      738 LBEKHAEOVSAOHLVHLKQGOHYBEKIKVLDNQIKDLADKETLENMQRHEEAAHE 797
Db      589 -----TEKDSVKRLQDEA-E 602
Qy      798 KGLILSEQKAMINAMDSKISLRQRYELSEANKLANSSLFTORMKKADEEMISELRQ 857
Db      603 KAL-----KQVADPEEKLEI-ETEKIAL-----IKKQEVITTEAR-- 637
Qy      858 KPYLETQAGLEAQRKLEEQLEKISHQDSHDKNRILETRLEVSLEHEEQKLELRQ 917
Db      638 -----KSEVTDHL-----SEEVVAAKNT 656
Qy      918 LTELQSLQRESQTLQAARALAESOLROAKTELETTAAEEIQTALTARHDIQRK 977
Db      657 IASLQATNERETEIKLK-----ORMDEERASHTAQSOEQKQLEAHYERAKM 706
Qy      978 FDLARRSCVITTLDEQNLQTEDMALNNQNFILSQDL--EASGANDETVQARSVDH 1035
Db      707 -----LQDVEQNNVENRGLRDEITELSQOMALPRGILNEQ--QIHEIFNW 751
Qy      1036 LRRE-ITEREMOLTSOKQT--MEALKTCTMLDEQVWDLBALNDELLEKORWEAMRSLV 1092
Db      752 VSEKATREEMENLTKITGEVSLKNNSPITTSNYI-----QNTBGMWS-- 797
Qy      1093 GDEKQPECEVRELQRM--LDTEKQSPARADQRTTESQVVELAVENKHAETIALQOALK 1150

```

```

Db      798 -----RRMNVARQDGLDQROLAETDAK-----LTKAKAEK 830
Qy      1151 EOKIKAESLSDKNDLEKKHAMLEMARSLQOKLETERELKQRLLEQAKLQOQMDIQ-- 1208
Db      831 NSQEQVITSAARLDDTEKMA-----SLMREVAAMLKQOKNIENS 869
Qy      1209 KNIIFRLTQGLQEAALDADLLKTERSDLEYQLENTQVLYSHEKYAMEGTISQOTKLIDFL 1268
Db      870 SDSAFSSTWG-----RGDLMTSMAND--YEMSN--SSLMOQEMISROSTSEYENAIL--- 917
Qy      1269 QAKMDQPAKKKKVPLQVNEILKLALEKAKCABLEEALQTRIELRARSAREAAHRAKATDH 1328
Db      918 -----LHDQVPPKRYDILRY--KQPM-----937
Qy      1329 PHSTPATARQOIAMSAIVSPHQPASMSLAPPSRRKSSSTPEEFSSRLKERMHNI 1388
Db      938 ----KTASGISPISISMERG--HNPFRMKIKTP-----966
Qy      1389 PHRPNGLANRATKICAVGLD-TYHFGQASKCLCEQVMCHPKSTCIPATCGLPAYATH 1447
Db      967 -----TKCGHCTISILIGLDRQGLFCQSCQYACHVSCAERYOSQCPVEE--- 1010
Qy      1448 FTBAFCRDKNNSPGLQTKERSSSLHEGMKVKVRNNRGGQGMQRKIVLEGSKVLLYDN 1507
Db      1011 -----ERRPLGMITR--AEQNMQOVKN 1031
Qy      1508 EAREAGQPVPEEFELCLPDGVSIGHAVGASELANTAKADVPYLK-----ESHPH 1560
Db      1032 EIR-----LVDMRDDEFTVCG--VSEADVIHQKGIPIKFRYTTQILNSSSEYS 1081
Qy      1561 TCPMGRTYLLADSPDKQWTALESVAVAGVSRKADAKLQNSLLKLEGDD-- 1617
Db      1082 SSKFYTLP-MAETEEKRKKVVALSELKT--LLRRSKLADR--AFVKEVPVVT 1133
Qy      1618 ----RLDMCTLPSPQVVLVGTBEGLYALNVLKNSLTHYPGIGAVFOYIYIKOLEKLM 1673
Db      1134 LPSIRVAQCCAILIDRSKIVIGFSDHGLYCIEIRQLLIPVG-----EK-- 1177
Qy      1674 IAGEERALCLVDYKVKVQSLAQSHLPAP-----DISPINEAVKCHLFG 1719
Db      1178 -ENKQRCVETVEYDEBAQOLMTVGPRAKHVRIVPSAALDGRLLKRIKVDYTGCHLLA 1236
Qy      1720 AGKIENG---LCICAMPKSVILRYNENLSKYCIRKEIETSESCIHFTNYSILGTN 1776
Db      1237 VGTNNPGRAGPAVPAFKSVTIFQIDRSEKHKWDLMPGTPGIAIFNGRLYVGS 1296
Qy      1777 KPYEI-----DMKQYTLF--EFLDKNDHSLAPAVPAASSNFPVSIQVNS 1820
Db      1297 HSPFRWSLVGVDSPVGSGDASGAVLQHISLVNMEDTSLQ---FLNQOTSVEAKLI-VNV 1352
Qy      1821 AGQREHYLCFHFEGVVDVSGRRSRPTDLKWSRLPLAFAYRBPYLPVTHPNSLEVIEI 1879
Db      1353 PGPSPDEYLVFNMTIGLVNEMGRSRLLPEVMFPQAKFYAHBYLTCFSENEVDIFNV 1411

RESULT 5
S74244
serine/threonine-specific protein kinase (EC 2.7.1.-) isoform I, Rho-associated - mouse
C/Spec: Mus musculus (house mouse)
C/Date: 29-Jan-1998 #sequence_revision 06-Feb-1998 #text_change 16-Aug-2004
C/Accession: S74244
R/Nakagawa, O.; Fujisawa, K.; Ishizaki, T.; Saito, Y.; Nakao, K.; Narumiyu, S.
FEBS Lett. 392, 189-193, 1996
A/Title: ROCK-I and ROCK-II, two isoforms of Rho-associated coiled-coil forming protein
A/Reference number: S74244; MIMD:96368048; PMID:8772201
A/Accession: S74244
A/Status: nucleic acid sequence not shown
A/Molecule type: mRNA
A/Residues: 1-1354 <NAK>
A/Cross-references: UNIPROT:P70335; EMBL:U58512; NID:g1514695; P1DN:AAC53132.1; P1D:g1514
C/Superfamily: protein kinase homology
C/Keywords: coiled coil; phosphotransferase; serine/threonine-specific protein kinase
P/74-338/Domain: protein kinase homology <KIN>

```



```

Db      67 NIDNFKLRVETKIVKIGLQKAKEDYDVYKVIIGKAGEVOLVHKHKSQKYYAKKLTSKF 126
Qy      132 ALTAQEVSPFEFEERNILSRSTSPWIPOLQYAFODKXHLYLWMEYOPGDLISLNRVED 191
Db      127 EMIKRSDSAFWEERDIMAFAFNSPMVYQLFCAFODDKLYLVMEYMGGDLVNLMSYND- 185
Qy      192 QLDENLIQFLAELILAVSHVHLMGYVHRDIKPEINILVDRTGHTKLYVDFGSAAMNSNM 251
Db      186 -VPEKMAKFFYTAEVVLLDALHSHWGLIHRDVKPDNMILLDKIGHKILADFGCMQIDETGM 244
Qy      252 VNAKLPICGTPDYMAPEVLTWVNGDGKGYGLDCDMWSGVYAYEMITGRSPFAEGTSART 311
Db      245 VHCDTAVGTIDYISPEVVKSGQGD- GYVGRCDMWSGVYLFEMLVGDTFFVADSLVGT 302
Qy      312 FNNIMNFORFLKFPDDPKVSSDFLDLIQSLICQOKERLKPEG--LCCHPFPSKIDNN-- 366
Db      303 YSKIMDKNSLCFPEDEAIEISGAKNLIICAPLTREYVLRNGVBEIKQHPFKNDQNMWD 362
Qy      367 NINSPPEPPTLKSDDDTSNDFEPEKNSM-VSSSPQSLSPSGSEBELPFVGFYSKAL 425
Db      363 NIRETAPVVPBELSSDSSNFDIEDDKGVETFPF--PKAVVGNQLPFIPTFYREN 419
Qy      426 GILGRSSVVGSLDSPA--KTSSMEKKLILKSKELOSODKCHMEQEMTRLHRVSEVE 483
Db      420 LLLS-----DSPCKENDSTQSR--KNESOEIQQKLYLLEHLS----- 457
Qy      484 AVLQKEVELKASETQSLLEQDLATYITECSLSKSLLEQARMVNSQEDDKALQLLDIR 543
Db      458 -----TEIQAKE-----LEQ-----KCKSVNTLEKAKALEEITLRKINVESTLR 499
Qy      544 --EQSRKLOEIKQEQYQAVQVEKRLMMQLEEDIVSARRSDLYESBLRESRLAAEEFKR 601
Db      500 QLRERKALLQHKNAEYORKADHEADKRNLENVDNYSIKDLQLEDKKNQNSQISTE---- 555
Qy      602 KARECHOKLLKADQCKREYGEVAKLEKIAEQOLKIOEIOE-----K 644
Db      556 KVNQLOQDLQETNALLTESDTPARLKTKQTESSKQIOLESNNRDLQDKNCLLETAKLK 615
Qy      645 LEKAVKASTATELLQNIROAKERARELEKLR-----EDSSEG--IRKLVAAEERRH 698
Db      616 LEKEP---INQVLSERBRDRHTGSEIINDLQIRIGLEEDVYNGKILAKVLEKQOL 672
Qy      699 SLENKVRLEMTERRERRLKDIQTSQOIQOMADKILLEEKREARQVSAOHLEVHLKQ 758
Db      673 Q-----ERFTDLKEKKNMEIDMTYQLKVIQOS-----LEQEBTEHKA-----TKARLAD 717
Qy      759 KEQHYE-----EKIVLNDQIKKDLADKETLENMQRHE-----EAAHEKGI 801
Db      718 KNRTYESIEBAKSEAMKEMEKULSEERTLKQKVENLLEAKKSCILDCDLKQSOQKINE 777
Qy      802 LSQOKAMINAMDSKIRLSIORIVLSEANKLAANSLSFTQRMKKAQEBMSILRQCKFYL 861
Db      778 LKQKQVLAN--EDVRNLTUKIEQETQKCL-----TQNDLMQIQOQVNTLMS----- 823
Qy      862 ETQAGKLEAQRKLEQLEKISHQSDSKNRLLLELTRLEVSLEHEQKLEKQKQTEL 921
Db      824 -----EKQKQK-----ENHLEMKKSL-----EKQNAELRKERQDA 855
Qy      922 QLELOERESULTALQAAARALBESOLROAKTELEETT--AAABEIQALTARDELQOKF 978
Db      856 DQGMKEIQDQLEAEQVSTLYKTOVRELKECECEKTKLQCKQKQELQDERSLAQL 915
Qy      979 DALRNSCTVITDLEEQNLQTEQNAELMNQNFYLSQOLDEASGANDSIVQLRBSVDHLR 1038
Db      916 E-----LTLTKADSEQLARSIAEQYSDLEKEKIMKE-----LEIKEMMARHK 959
Qy      1039 EITEREMQLTQKQYQMEALKTCTCMLEEQVMDLEALNDELLEKRCQWEARSVLDGE--- 1095
Db      960 ELTEKQATIASLETBTRTTLTSVAVLANL---EKELNNKLEKAQEQ---LSRLKBEIS 1012
Qy      1096 ----KQOFERQVRELQRMLDTEKQSRARADQRTTESQVAVLAVKEKAEILALQALKE 1151

```

```

Db      1013 AAATKAQE-----KQLTE 1027
Qy      1152 QKLKASLSDKDNLDEKKHALLKEMNARSLOQKLETERELKQRLLEBQAKLOQOMDLOKMH 1211
Db      1028 RLTKQAVNNKLAEIMNRKEPVKRGNDPVRKEKENRKLHMLSEREKLTQOM----- 1081
Qy      1212 IFRLLQLEALDADLLKTERSDLEQLENTIQVLYSHEKVKMEGTISQOTKLIDFLQAK 1271
Db      1082 -----KQYQELNEMQAIABESQI----- 1101
Qy      1272 MDQPAKKKKVPLQYNEKLKALEKEKARCALAEALQKTRIELRSAREBAARKATDHPH 1331
Db      1102 -----RIEQMTLDSKOSDIEQLRSQALHIGLDS----- 1133
Qy      1332 STRATARQOIAMSAIVSPHOPASMLAPPSRRKSSSTPEEFSSRLKERMHNIPHR 1391
Db      1134 -----SIGSGRGDTAADGFPES----- 1151
Qy      1392 FNVGLNRAATKCAVCLDTVHFGROASKCEQVMCHPKCSTCLPATGCLPAEYATHFTEA 1451
Db      1152 ----- 1151
Qy      1452 FCRDKKNSPGLQTEPSSSLHSGMKVPPRNKRQGGQGMKRYIVLSSSKYLYDNEARE 1511
Db      1152 -----RLBGMISLPVRNNTYKKGWAKYIVSSKILFYDSEQDK 1191
Qy      1512 AGORP 1516
Db      1192 EQSNP 1196

```

```

RESULT 7
S69211
serine/threonine-specific protein kinase (EC 2.7.1.-), Rho-associated - human
C:Species: Homo sapiens (hmn)
C>Date: 06-Dec-1996 #sequence_revision 06-Dec-1996 #text_change 16-Aug-2004
C/Accession: S69211, S71910
R:Ichikaki, T.; Maekawa, M.; Fujisawa, K.; Okawa, K.; Iwamatsu, A.; Watanabe
EMBO J. 15, 1885-1893, 1996
A>Title: The small GTP-binding protein Rho binds to and activates a 160 kDa Ser/Thr prot
A/Reference number: S69211, MIMD:96203110, PMID:8617235
A/Accession: S69211
A>Status: nucleic acid sequence not shown
A/Molecule type: mRNA
A/Residues: 1-1354 <ISH1>
A/Cross-reference: UNIPROT:Q13464; EMBL:U043195; NID:G1276900; PTDN:AB02814.1; PTD:G127
A/Accession: S71910
A/Molecule type: protein
A/Residues: 187-195;281-288;465-473;573-587;818-828;885-893;934-945 <ISH2>
C/Superfamily: protein kinase C zinc-binding repeat homology; protein kinase
C/Keywords: ATP; phosphotransferase; serine/threonine-specific protein kinase
F:82-90/Region: protein kinase ATP-binding motif
F:1229-1287/Domain: protein kinase C zinc-binding repeat homology <KZN>

```

```

Query Match      12.0%; Score 1260; DB 2; Length 1354;
Best Local Similarity 26.3%; Pred. No. 5,9e-27;
Matches 400; Conservative 293; Mismatches 531; Indels 296; Gaps 53;

Qy      44 SPLSRGILDALFVLEECGQPALMKIKHVSNFVRKYSDTIABLQELQPSAKQFEVSLV 103
Db      23 SEVNSDCLDLGLALVYLDLPALRKKNIDNLFSLRYKQDTINKRIDRMKRAEDVEYVKVI 82

Qy      104 GCGHFAEVQVVRKAKATDIYAMVMKKKALLAQEQVSFFEEERNILSRSTSPWIPOLQYA 163
Db      83 GRAFGEVQVLRHKSTRKYVAMKLISKFEMIKSDSAFFWEERDIMAFAFNSPMVYQLFYA 142

Qy      164 FQDKNHLYLWMEYOPGDDLISLNRVEDQDENLIIQFLAELILAVSHVHLMGYVHRDIK 223
Db      143 FQDDRYLYVMMEYMPGGDLVNLMSYND--VPEKMARVYTAEVVLLDALHSHWGLIHRDVK 200

Qy      224 PENILVDRGTHTKLYVDFGSAAKNSNMKNVNAKLPIGTPTYMAPEVLTWVNGDGKGYGLD 283

```

```

Db 201 PDNMLDKSHLKIADBGTCMKRKNKGMVRCDTAVGFPDYISPEVYLKSGGD--GYTGRE 258
Qy 284 CDWMSVGVIAIYEMLYGRSPAEAGTSARTFNINNFQFLFKPDDPKVSDFLDIQSLIC 343
Db 259 CDWMSVGVFLYEMLVGDTPFYADSLVGTYSKINHKNLSLTFPPDNDISKEAKNLICAFLT 318
Qy 344 GQXERLKFEG--LCCHPFPSKID--WNINRSPPPVPTLKSDDTSNDFEPKSNVWS 398
Db 319 DREVRLEARNVBEIKRHLFFKNDQMAWETLRDTAVAPVPLSSDIDTSNDDLEEDGEE 378
Qy 399 SS-PCOLSPSGSGEELPFVGFSSYKALGILGRSESVSGLDSPAKTSMEKLLIKSKE 457
Db 379 EITPTI---PKAFVGNQLPFVGFTI-----YSNRRTYSANPNDRNTSSN-----ADKS 423
Qy 458 LODS-QDKCHMEQEMTRLHRRVSEVAVLQKEVELKASGTOR---SLLEOD-LATYIT 512
Db 424 LQSLQTIYIKLEBQL-----HNEMQKDEMEQCRISNLIKDKIMELD 468
Qy 513 ECGSLKRSLEQARMEVSGEDDKALQILHDTREOSRKLQETKEQYQAOVE-ENRLLMNOL 571
Db 469 EENGNORNLST--VSQIEKEKMLQHRINEYORKAQ--ENKRNVEVEVSTLKDOL 523
Qy 572 EBDLVASRRSDLYESELRESRLAEEFKKATKCOHKLKAKQOGKEVGEVAKLEKIN 631
Db 524 -EDLKVSQNSQL-----ANF--KLSQLOQLBANDLRTESDTRAVRLKSH 568
Qy 632 AEODLKIOELOEKLEKAVKASTATELLONIRQAKERARELEKQNRDESSSEGIRKLV 691
Db 569 TENSKTSISQL-ESLNRELQ--EENRLTEN--SKQTDMDYVQL-----AIL 610
Qy 692 EAEBRHSLENNKRYLETMERENRLKODIQTK---SQOIQOMADKILEEKHREAQV 747
Db 611 EAERBRDQ-----HDESEMIGDQARTISLOEEVKHKLKHLMEKVEGERKEAOD 657
Qy 748 SAQHLVHLKQXQGHYEKIKVLDNQKKDLADETLENNMQHREBAHE---KGRILS 803
Db 658 MLNHSE---EK-----NNLEIDLNRY--LKSLOQRLEQVNEHVKTKALTD 700
Qy 804 EOKAMINAMDSKIRSLFORIVELSEANKLANSSLFTORMKAOEENIS-ELRQOKVYLE 862
Db 701 KHSISIEAKSVAMCEMKIKLEKEAREKXENRYV---QIEQCSMLVDVLKOSQOKLE 756
Qy 863 TOAGKLEAONRKLEEOLEKISHODHSDKNRILLEITREVSLEHEOK----- 911
Db 757 HLGNKX-----RMDEVTGNLTLOEGSNKRLLONELKTOAFRADMLKGLKQMOEIN 812
Qy 912 --LEIKROLTELQLS-----LOERESQTLTQAARALBESQLAQATELEET 957
Db 813 TLLEAKR-LLEFELAQLTKQYRGNEGQMRLEODLEAOYFSTLYKTVKXLEKEIEKN 871
Qy 958 AAABEELIOTALTAHDEIORKFDALRNSCTVITDLEBQNLQTEDNAELNNQNFYLSKOLD 1017
Db 872 RBNLKTKIOELONEKETLATOLD-----LAETKASEQLA--KGLLEQYFELTQESK 921
Qy 1018 EASGANDEIVOLSEVDHLRREITEREMOULTSQOTMEALKTCTMLBEOVMLEALNDE 1077
Db 922 KAASRN-----KQETIDKD-----HTVSRLSEANSMLTQDIEILRRENEE 961
Qy 1078 LLEKERWEAMRSVLGDE-----KSQFECVRELQRLMDTEKOSRADAQRITESQYVE 1132
Db 962 LTERMKAAEEYKLEKEEISNLKAAFE-----KNINTERLTKTQAVNKLAEIMNRKD 1014
Qy 1133 LAVAKHKAELIALQOALEKQK--LKAESLSDKLNDEKHAMLENNARSLOQCLFETERE 1189
Db 1015 FKIDRKKAANTQDLRKKEKKNRKQLELNOERKFNQWVXK-----QKEL---ND 1061
Qy 1190 LKORLLEEOA---KLQOQMDLQKNHIFRLTQGLQALDRADLL-----KTERSDLEYOL 1240
Db 1062 MQQOLVBEACHRNELQWOLASKSDIEQLRAKLLIDLSSTSVASFPSSADEFDGNLPESRI 1121
Qy 1241 ENIQVLSHKKVMEGTISQ---QTKLIDPLQAKMDQPAKAKKVPPLQYNELKALAEKX 1296
Db 1122 EGMVLSPNRGNIKRYGKWKQYVSVSSKILLFYNDQDQKEOSNPSMVIDIKL----- 1173

```

```

Qy 1297 ARCAELEALQTRIELRSAREEAAHKAATDHPHPSTPATARQOIAMSAIVRSEHPSA 1356
Db 1174 -----FHVRRPTQGDVYRAET-----EELPKI 1195
Qy 1357 MSLLAPSSRRKSSSTPEESRRLKERMHNIIPHRFVGNLMBRATKCAVCLDVHFGROA 1416
Db 1196 FQILYANEGECRQDVEHPEVQAEKTNFOHKGHEFLPTLYHPRANCADAKPLMHVFKP 1255
Qy 1417 SKCLCEQVMCHPRCS-----TCLPATCGIPAEVATHTPEAFCRDRNNSPGIQTKE 1466
Db 1256 PRLAEGR-RCHVAGCHDHLDKEDLIC-----PCQVSYDVSTA--RDWLLILACSGQDEQ 1305
Qy 1467 PSSSLHLEGMMKVPNNKRG 1486
Db 1306 KKWVTHLV--KKIPKNPPSG 1323

RESULT 8
S74245
serine/threonine-specific protein kinase (EC 2.7.1.-) isoform II, Rho-associated - mouse
C:Species: Mus musculus (house mouse)
C>Date: 29-Jan-1998 #sequence_revision 06-Feb-1998 #text_change 16-Aug-2004
C:Accession: S74245
R:Nakagawa, O.; Fujisawa, K.; Ishizaki, T.; Saito, Y.; Nakao, K.; Narumiya, S.
FEBS Lett. 392, 189-193, 1996
A:Title: ROCK-I and ROCK-II, two isoforms of Rho-associated coiled-coil forming protein
A:Reference number: S74244; WUID:96368048; PMID:8772201
A:Accession: S74245
A:Status: preliminary; nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-1368 <NA>
A:Cross-references: UNIPROT:P70336; EMBL:U58513; NID:G1514697; PID:AA053133.1; PID:G151.
A:Superfamily: protein kinase C zinc-binding repeat homology; protein kinase homology
F:Keywords: phosphotransferase; serine/threonine-specific protein kinase
F:90-354/Domains: protein kinase homology <KIN>
F:161-1315/Domains: protein kinase C zinc-binding repeat homology <KZN>

Query Match 11.9%; Score 1250.5; DB 2; Length 1388;
Best Local Similarity 24.1%; Pred. No. 1,1e-26;
Matches 365; Conservative 284; Mismatches 489; Indels 379; Gaps 35;

Qy 32 QGKPPPTQOQMSPLSBEGLDLYFVLFEECSOPALMKIKNVSNFPAKYSDDTAELOELQ 91
Db 27 QRLKELAIRDRPRSPINWESLIDGLNSLVLDLDPALRKNNKINDPFLNRYEKIYKJRLQ 86
Qy 92 PSAKDFEYNSLVCGGFAEVQVVRKATGDIYAMKWKKKALLAOEVSFEEERNITSR 151
Db 87 MKAEVDVAVVIGRGAEGEVQVRHRAISOQRYAMKLSKFEEMTKRSDSAFWEERDIMA 146
Qy 152 STSPWIPOLQYAFQDKNHLYLVMEYOPGGLDLSLNRVEDLDENLLOFYAEILIAVHS 211
Db 147 ANSPWVQVLFCAQODRILYLVMEYMPGGDLVNLMSNYD--VEKNAKFTAAVLAIDA 204
Qy 212 VHLMGVHRDIKENILVDRTHIKLVDPGSAAMNSKNVNAKLPIGTVDYAAPEVLTIV 271
Db 205 IHSMGILHRDVKEDNMLDKHGHLKIADFGTCMKMETGWHVCDTAVGTPDYISPEVLKS 264
Qy 272 MNDDGKGTVGLDDCMDSVGVIAIYEMLYGRSPAEAGTSARTFNINNFQFLFKPDDPKV 331
Db 265 QGGD--GYTGRECDWMSVGVFLYEMLVGDTPFYADSLVGTYSKINHKNLSLTFPPDNDIS 322
Qy 332 SDFLDLIQSLQCGQERLKFEG--LCCHPFPSKIDWN--NINRSPPPVPTLKSDDTS 386
Db 323 KXAKNLICALTLTREVRLGNGVBEIKQHPFFKNDQNMNMNINETAAPVPELSDIDS 382
Qy 387 NFDEPEKNSW-VASSPQASPSGSGEELPFVGFSSYKALGILGRSESVSGLDSP--AK 443
Db 383 NFDDIEDDKGDTVETPLI--PKAFVGNQLPFIFTYFRRENILIS-----DSPPCRE 430
Qy 444 TSMMEKLLIKSELDSDQCHKRMEQMTRLHRRVSEVAVLQKEVELKASGTORSL 503
Db 431 NDAIQTR--KSESEOIQKLYALEHLS-----SEVOA--KELEOK----- 469

```

```

QY 504 EODLATYITBSSSLKSLGQARMESVQEDDKALQJLHDIR--EGRKLTQITKEQYQAV 561
Db 470 -----CKSINTREKTAKELEBEITLRKVESTTLQLEBEKALLQHKNAEYORKA 519
QY 562 EEMRLMNOLEEDIVSARRSDLYESSELRESRLAAEFKFRATSCQHLTKAQOGKEV 621
Db 520 DHEADKGRNLENDVNSLKQJLEDKKGNQSQITE---KVNQJQJLDBANMLLTRES 575
QY 622 GEYAKLEKINAEOQLKQELQEKLEKAVKASTATELLQNIROKERAERELEK-LONRE 680
Db 576 DTAARLKRTOAESKQIQOLESN-----NRDLQJNGCLLETAKLKLEKEFINLQ 624
QY 681 DSSGIRKKLVAEERHSLNENKTKLETMERRENRLKDDIQTSQQLQOMADKILEEE 740
Db 625 SALESERDRTHSGEIIINDQGRSLEEDLKTGKALLAKELERQJQ--EKLTDLEK 681
QY 741 KHRQAQVSAQH---LEVHLKQKE-QHYEEKIKYLD-NOIKQD--ADKETLEMORHE 792
Db 682 EKSMEIDMTYQKLVIOQSLEQEAHEKTTKARLADNKKIYESIEEAKSEAMKEMEKLL 741
QY 793 EEAHEKQKI--LSEQRAMINAMDSKJRSLEQRIVELSEANKLAANSFLTQRMKAQEE 849
Db 742 EERSLKQKVENLLEAEKSGSILDDCLKQSQKLNEL-----LKQXDV 784
QY 850 MISLRQOKTYLETQAGKLEAQNRKLEBQLEKISHQHSKNRLLEETRLREVSLEHEE 909
Db 785 LNEBVRNLTLEIBETQRCIMQNDLTKQITQOVNTLKMSEK-QIKQENNHMEKMKYLEK 843
QY 910 QKLELQKQLTQJLSLOERESQTLQALQARALESQJLQAKTELEPT--AAEMEEQLA 966
Db 844 QNTLEKREKQADQOMELQDQLEAEQVFTLYTVARELEKEEBEKKLCKELQJQKOD 903
QY 967 LTAARDEIQRKFDALRNSCTVITDLEBQNLQTEDNAELNNQNFYLSKQJLEASGANDI 1026
Db 904 LQDRBDSLQJLE-----ITLTKADSEQLARSIAEQYSDELEKIKME----- 947
QY 1027 VQLRSEVDHLRREITEREMQJTSQKQTEALKTTCYMLEBQVMDLEALNDELLEKERQWE 1086
Db 948 LEIEMMARHQEILTEKQJTLASLEETNRILTSDVANLAN--EKEELNNLXKQSQJQ-- 1002
QY 1087 AMSVYLQDE-----KSQFCRVRLEQRLMDTEKQSRARADQRTESRQVVELAVENK 1139
Db 1003 --LSKDEEMSAANAIAQFE----- 1021
QY 1140 AEIATLQALKEQKTKAESLSDKLNDEKKAULEMNAARSJQKLETERBELQJLEBOA 1199
Db 1022 -----KQJLNERLTQAVNKLAEIMNRKEPVKSGSDTVRRKEKENRKLHMLKSERE 1075
QY 1200 KLOQOMDLOQNHILFRLTQJLOEALRADLLKTERSDELEYOLENIQVLYSHEKVMEGTIS 1259
Db 1076 KLTQOM----- 1096
QY 1260 QQRTKLIDFLQAKDQPAKKKKVPLQVNELKLALEKEKARCALEBALQKTRILERSAREE 1319
Db 1097 EESQI-----RIELQJTLDSKQSDIEQLSQCJALHIGMDS-- 1133
QY 1320 AAHRKATDHPSTPATARQOIAMSAIVRSPHQPSAMSLAPPSRKSSSTPEERSR 1379
Db 1134 -----SISGPGDABPDGFPES----- 1151
QY 1380 LKERMHENIPIHRFNVGLNMPATKACVLDTVHFGQAQSKLEQJWCHPKQSTCLPATCG 1439
Db 1152 ----- 1151
QY 1440 LPAPAYATHTFPAFCRDMNSBGLQTKPESSSLHLEGMMKVPNNKQJQOGDKRYITYLEG 1499
Db 1152 -----RLLEGWLSLPIVRNNTKKFGVWKYIVVSS 1179
QY 1500 SKVLYIDNEARBAQGRP 1516
Db 1180 KKLIFYSQDQKQSNP 1196

```

```

RESULT 9
T25539
hypochemical protein C10H11.9 - Caenorhabditis elegans
C/Species: Caenorhabditis elegans
C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C/Accession: T25539
R/Dante, M.; Wambale, P.
submitted to the EMBL Data Library, February 1997
A/Description: The sequence of C. elegans cosmid C10H11.
A/Accession number: Z20047
A/Accession: T25539
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-1173 <DAN>
A/Cross-references: UNIPROT:P92199; EMBL:U88311; P1DN:AA842348.1; GSPDB:GN00019; CESP:C1
A/Experimental source: strain Bristol N2; clone C10H11
C/Genetics:
A/Gene: CESP:C10H11.9
A/Map position: 1
A/Intons: 23/3; 51/1; 104/3; 343/1; 478/2; 868/3; 891/3; 970/3; 1027/3; 1114/1
C/Superfamily: hypochemical protein C10H11.9; protein kinase homology

```

Query Match 10.7%; Score 1126.5; DB 2; Length 1173;

Best Local Similarity 29.2%; Pred. No. 1.8e-23; Matches 294; Conservative 198; Mismatches 339; Indels 175; Gaps 30;

```

QY 44 SPSRBEGLDALFVLEFECQPALMKIKHVSNEFKYSDTIAELOEPSAKDFEVSILV 103
Db 15 SPINIESLDTITNLVNDCKIPVLMRKYSVDNFSRERVESLAALRMQADRLQKVI 74
QY 104 GCGFAEVOVREKATGDIYAMKMKKALLAQVSPFEERNILSRSTPMIPLQJYA 163
Db 75 GRGAFGVHLVRRRTNTVYAMKMLNDKMDIKRADSAFFWEERIDIMAHANSEWIVRLQYA 134
QY 164 FQDKNHLIYMEVQPGDLSLNKRYEDQJDENLQCYTLAEILAVSVHMGVNHDIK 223
Db 135 FQDPRLHYMEVYPPGDLVNLMTSYE--VSEKTRFYTAIEYALALHSMGVIHDDVK 192
QY 224 PENLVDRTGHILKLVDSAAKNSNKNVNAKLPIGPPDVAAPVLTVMNGDGKGYGLD 283
Db 193 PDNKLISIGHIKLADGTCVKNANGVRCSTRAGPDIISPEVLNQGODAE--FGKE 250
QY 284 CDWWSGVIAEMILYGRSPFAEGTSARTFNINMFQRLKFPDPKVSDFLLDILQSLC 343
Db 251 VDMWSGVFIYEMLVGEPYAEALVSTYTNIMHKSLKFPDPEPLISTQAKDIIKKFLS 310
QY 344 GQKRL--KFEGLCCHPFFSKIDMN--NIRNSPPVPVPLKSDDDTSNDEPKNSWVS 398
Db 311 AADPRLGRNSVDDIRNKKFFVNDMEWTPATLREASPPVPIELKSDDDTTFHELETRDRDN 370
QY 399 SSPQQLSPSGFSGELPFGVFSYSKALGILGRSESVSGLDSPAKTSSMEK----- 449
Db 371 AGDPQL-PKTFNGNLFPFIQFTYSNEYSV--KNLKGAGAGSKQGIQDHKQJTVBQ 426
QY 450 -----KLLIKSKB-----LQDS-----ODKCHQEQEMTRLRRAVSEV 482
Db 427 PLTNHASGVPEEKYEAANKELDSKREFELNDSIARNIRIAPKMINENKNSLSTKISDL 486
QY 483 E-----AVLSQKEVLEKASETQRSGLBEDLATYITECSLKSLSQAR 525
Db 487 ERELKONKDLRAGADSDAKYNEIAVELRMSKENSMESELKFRDKCEQLKEDLRKKS 546
QY 526 MEVSOEDDKALQJLHDIRQSRKLOEIKQDEYQAVQVEKMLMNOQJLEEDIVSARRSDLY 585
Db 547 GEIQAQENETQVFOQKQDDEAPAEIKRDYELLQIRENKSV-QLKAL-DEKKEGAY 604
QY 586 ESELRESRLAAEFKFRKATECOHKLKAKQOGKEVEYAKLEKINAEOQLKQJLEOKL 645
Db 605 QQSVAKATDA--EMERMGVFEKQJLEHANDERKRE-----EQKRTAAEFQGR 650
QY 646 EKAVKASTAET-ELLQNIROAKERAERELEK-LQNRDSSBGIRKCLVABERRHSLNENK 703

```

```

Db      651 VAKTACIENAYELQNDYKSMKEARDLBRDIOD-----VITEKR----- 692
Qy      704 VKLEIETWERENRKIDIQTKSQOIQOMADKIIIELEE--KREAOVSA-----QHLEVHL 756
Db      693 -----LEIRVEQIQMDSRNTERVLTSLQDEIVASQGEAKYKEDGLKIDGFXGHELENE 746
Qy      757 KQKQNHHEEKIKYLDNOKIKDLADKE-----TLENNQORHEEHAHKIKLSQKXMIN 810
Db      747 KMTQITLLENILV-----ADKERGMLKMEVQELMQRHKWITNK----- 785
Qy      811 AMSDKIRLEQRIYELSEANKLAANSILFTORNNKAQOEMISELROQFYLETQAGKLEA 870
Db      786 --DQTLKHLTQDDEIRQOQKIESSEQ---ESNDK---QITADIRK--LDLEKAKHKA 834
Qy      871 QNRKLEBQLEKISHQDSD-----KNRLLEETRLREVSLSHEEQL-----ELKQ 917
Db      835 VINKLEBMAKROPLKKGKGVTKSALIKERREI--MALRQEDTWKRIATLAFYENDKQ 892
Qy      918 LTELQSLQRESQTLTQAKRALLESQLRQAKTELETTAAEE 963
Db      893 AEHNIAIOD-----MQTTQDALRDLKCKEKEELANRVNTRYE 931

```

RESULT 10

B49364 protein kinase (EC 2.7.1.37), myotonic dystrophy-associated - human

N;Alternate names: myotonin protein kinase

C;Species: Homo sapiens (man)

C;Date: 20-Feb-1995 #sequence revision 20-Feb-1995 #text_change 16-Aug-2004

C;Accession: B49364; S46881; A42101

R;Shaw, D.J.; McCurrach, W.; Rundle, S.A.; Harley, H.G.; Crow, S.R.; Sohn, R.; Thirion, Genomics 18, 673-679, 1993

A;Title: Genomic organization and transcriptional units at the myotonic dystrophy locus.

A;Reference number: A49364; MUID:94140369; PMID:7905855

A;Accession: B49364

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-624 <STA>

A;Cross-references: UNIPROT:Q09013; GB:LI9268; NID:G3071176; PIDN:AAA36206.1; PID:G3071177

R;Sasagawa, N.; Sorliachi, H.; Maruyama, K.; Arahata, K.; Ishiura, S.; Suzuki, K. FEBS Lett. 351, 22-26, 1994

A;Title: Expression of a novel human myotonin protein kinase (MCPK) cDNA clone which enc

A;Reference number: S46881; MUID:94357271; PMID:8076686

A;Accession: S46881

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-417, 'I', 419-468, 'P', 470-544, 'MAPRPMASARWGOAPCTATCCSLRS', 'GLAVRRRPPCS

A;Cross-references: GB:S72883; NID:G633864; PIDN:AB31800.1; PID:G633865

R;Brook, J.D.; McCurrach, M.E.; Harley, H.G.; Buckler, A.J.; Church, D.; Aburatani, H.;

S.; Davies, J.; Shelpourne, P.; Buxton, J.; Jones, C.; Uuvonen, V.; Johnson, K.; Harper, Cell 68, 799-808, 1992

A;Title: Molecular basis of myotonic dystrophy: expansion of a trinucleotide (CTG) repe

A;Reference number: A42101; MUID:92154692; PMID:1110900

A;Accession: A42101

A;Molecule type: mRNA

A;Residues: 'PPQVQRGLAVGGA', 57-550, 'WLWASARWGOA', 564-624 <BRO>

A;Cross-references: GB:M914203; GB:M91465; NID:G186755; PIDN:AAA64884.1; PID:G186756

A;Note: sequence extracted from NCBI backbone (NCBI:82640, NCBI:P:82641)

C;Genetics:

A;Gene: GDB:DM

A;Cross-references: GDB:119097; OMIM:160900

A;Map position: 19q13.2-19q13.3

C;Superfamily: protein kinase homology

C;Keywords: ATP; phosphotransferase

F;69-339/Domain: protein kinase homology <KIN>

F;77-85/Region: protein kinase ATP-binding motif

```

Query Match      8.5%; Score 888.5; DB 2; Length 624;
Best Local Similarity 37.9%; Pred. No. 2.2e-17;
Matches 196; Conservative 103; Mismatches 177; Indels 41; Gaps 11;
Qy      46 LSRGGIIDLFLVFECSQPALMKIKIVSNFVKYSPTTIAELOLQPSADFEVRLVGC 105
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

```

Db      20 LGLPFLDLLLGVHQLGASLDAQKYVADFLQWAEPIVVRLEKVRLOQDDFEILKVIGR 79
Qy      106 GFAEYQVYREKATGDIYAMKVMKKALLAQEYVSFPEERNLISSTSWIQLQYAFQ 165
Db      80 GAFSEVAVVMQKQGVYAKKINKWMDLKRGVSCFRRERDVLVNGDRWITQLHFAFQ 139
Qy      166 DKRHLVLMVEYQGGDLTLNREYQDLNENLQFIATLIELILAVSHVLMGVYHARDIKPE 225
Db      140 DENVYLVMEYVYGGDLTLTLSTFGERRIPAMARFLIAELVMAIDSVHRLGYHARDIKPD 139
Qy      226 NILVDRTGHIKLVDFGSAAMKMSKNVNAKLPIGTPDYNAPEVL-TVMNGDKGYGLDC 284
Db      200 NILDDCGHRLRADFGSCJLRADGTVRSLVAVGTEDYLSPEILQVGGGPGTGYGPEC 259
Qy      285 DMSVGVYAVEMTYGSSPPAEGTSARTFNINMFOFLKRP--DDPKYSSFLLQSLIC 343
Db      260 DMVALGVAVEMTYGQTPFPAQDSTAETGYGVYHKKLSPLVDEGPEBARDFIQLLC 319
Qy      344 GQKERLKFES---LCHPFSKIDMNNINSPFPVPTLKSDDDTSNFDEPEKMSWSSS 400
Db      320 PPETRLGRGAGDPFRTHPFFFGLDWDGLRDSVPFTPDREGATDTCNFDLVEDGLTAMET 379
Qy      401 PCQLSPSGFSGEELPFVGFYSKALGILGRSESVSGLDSPAKTSSWEKKLT--IKSKE 457
Db      380 LSPDRGAPLGVHLPPVGYISY----CMALRDEVPG---PTMEVEABQLLEPHQAPS 432
Qy      458 LQDS---QDKCHKM-----EQEMTRLHRSVEVAIVLSQEV--ELKASTORS 501
Db      433 LEPVSVPQDETAIVAVPAAPAAEAELVRLQALREBEVLTROSLSREMAIRTD-- 490
Qy      502 LLEQDATYTTCESSILKRSLE-----QARMEVSQSD 532
Db      491 --NQNPASQUREAEARNRDLAEHVROLQERMEILOAE 525

```

RESULT 11

S71829 serine/threonine-specific protein kinase (EC 2.7.1.-) - mouse

N;Alternate names: myotonic dystrophy-associated protein kinase; myotonin protein kinase

C;Species: Mus musculus (house mouse)

C;Date: 06-Dec-1996 #sequence_revision 06-Dec-1996 #text_change 16-Aug-2004

A;Accession: S71829; S38815

R;Mahadevan, M.S.; Amemiya, C.; Jansen, G.; Sabourin, L.; Baird, S.; Neville, C.E.; Worme

submitted to the EMBL Data Library, January 1993

A;Description: Structure and genomic sequence of the myotonic dystrophy kinase (DM kinase)

A;Reference number: S71829

A;Accession: S71829

A;Molecule type: DNA

A;Residues: 1-557 <MAH>

A;Cross-references: UNIPROT:P54265; EMBL:Z21505

R;Mahadevan, M.S.; Amemiya, C.; Jansen, G.; Sabourin, L.; Baird, S.; Neville, C.E.; Worme

Hum. Mol. Genet. 2, 299-304, 1993

A;Title: Structure and genomic sequence of the myotonic dystrophy (DM kinase) gene.

A;Reference number: S38815; MUID:93271990; PMID:8499920

A;Accession: S38815

A;Molecule type: DNA

A;Residues: 1-53 <MAH>

A;Cross-references: EMBL:Z21503

C;Genetics:

A;Intons: 54/1; 84/3; 112/3; 144/3; 194/2; 225/3; 294/3; 382/3; 411/2; 448/3; 503/2

C;Superfamily: protein kinase homology

C;Keywords: ATP; phosphotransferase; protein kinase

F;69-339/Domain: protein kinase homology <KIN>

F;77-85/Region: protein kinase ATP-binding motif

```

Query Match      8.3%; Score 872; DB 2; Length 557;
Best Local Similarity 36.3%; Pred. No. 5.4e-17;
Matches 201; Conservative 103; Mismatches 190; Indels 50; Gaps 13;
Qy      46 LSRGGIIDLFLVFECSQPALMKIKIVSNFVKYSPTTIAELOLQPSADFEVRLVGC 105
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      20 LGLPFLDLLLGVHQLGASLDAQKYVADFLQWAEPIVVRLEKVRLOQDDFEILKVIGR 79

```


[illegible]

Db	231	FHRTHYNWPSTLPPDFISKPFESKRKAETWKRRBRAVAYSTVGTPDYIAEVEFO--PNG	289
Qy	275	DOKGYGIDCOWMSVGLAYEMIVSRAPEBTSARTFNINNNFOFLKPEPPDPKXSSPF	334
Db	290	-----YTSCDWMSGVIMTEMLIGYPPPCSELPOETRYKVIWMQOTLVEFSDVPIISA	344
Qy	335	LDLIOSLTCQEKERLKEFEG-----LCGHPFESKIDMNNIRASPPFVPTLSDDTSNF	388
Db	345	KATIRPFCEAEERRIGNNGDLEIQOC--PFXKRIDMNHIRERPPPIRTVVSIDTSNF	402
Qy	369	DE-PEKN-SWSSSPCOUSPSGFSSEELPEVGFYSK-----	423
Db	403	DDPFEDTLWPTST--LIRPEQOPRGGEFVDFYKREFDGLTQKMYILGWLFGEMLWVI	460
Qy	424	-----ALGILGRS-----BEVVG---LDSPATSSMEKKLLIKSKEL	458
Db	461	LSWGLVAVLCYFALNHLISGTGEAHTSTFGSVSGVANGVANGTHTSSGE-----WSNFI	515
Qy	459	QD-SODCKHMEQEMTRLRHRAVSEVEAVLSQCEVELEKASETORSLEYODLATYITECSL	517
Db	516	VSMLYSNHKVPAPMDAMIKSLNE-----AAKV--TSPTKCEVLFEBFGSHDYLQRP	567
Qy	518	KRSLEQAMEVVSQEDDKALQLLHDIRESRKLQEIKEOYQAVEBMRIMNNQLEBDVVS	577
Db	568	K--ISNVAVEHQPRD-----HLTVRSNVHLPCVKLRVSSQRPERRIVSNYDVSIYD	618
Qy	578	ARRRBDLYESELRESERLAAEFKFKRATCOHKLTKAKOQRPVEVEYAKLEKINAEQOLK	637
Db	619	LRG-----EVECRMACIANQLY-----LMGCFPSGREMEMIELRNDPNAOYVS	662
Qy	638	IOELOEKLEKAVKASTEATELLONIROAKERAERELKLONRDESEGIRKVLVAEERR	697
Db	663	IPMBESIRCLLSA-----VTINISEMPE-----BRNTPADTFSRTIYNPPIHS	709
Qy	698	HSELENKVRLEETMERENRLKDDIQTKSQOIQOMADKILLEBKREAOVSAOHLVHLK	757
Db	710	PS-----GTMRSYNNYDNTLHVDVAGTQ-----SSHRSNSPGSDVEMPEFK	751
Qy	758	QKEQHYEKIKIVLNOIQKDLADKETLENMQRHEEBAHEKILISEQKAMIMAMDSTR	817
Db	752	LNESH-----LISPSYNTNLSIPNKR-----LISPSYNTNLSIPNKR	772
Qy	818	SLEGRIVELSEANKLAANSLEFTQRMKAQOEEMISELKQCFYLETQAGKLEAQRKLEE	877
Db	773	-----IKVITQSRHGREVS-----QPFV-----NVEWDE	796
Qy	878	QLEKISHODSDKORLLELETRLREVSLEHEBOKYLEKROUTTELQLSQERESQUTALQA	937
Db	797	PAOKYSTTKGINAPFWE-----ETPFDDLTPATEIILFEIYEGNDKRLHMND	844
Qy	938	ARALLESOLROAKTELEETTAEBEIOALTANHDELO---KRPDALRNSCTYITDLEE	993
Db	845	-----EGFLGIAIVNFEERISGE-----TVHSLKIQGRPYRK-DATSGEITVOFDFFY	892
Qy	994	QLNQLTEDNAELNONFYLSKOLBEASGANDIEIOLUSEVNHLLREITEREMOULTSOQT	1055
Db	893	DPNLLTS-----GKLTPTVAKTNGSEBFRETLTTHRRPIYIDPHDNPGEHPI	940
Qy	1054	MEALKTTCTMLEOVMDLEALNDELLEKEROEMAMRSVLYGDEKQSEFCRVELEQRMIDTE	1133
Db	941	VPSKTTVTYV-----VKTVSQTLKEK-----PTIOSVHG-----SLENAVDPBA	977
Qy	1114	KQSAARADQRTTESROVVELAVKHKAEIILAOALKEQKJAKASLSKJLNDLEKHAML	1173
Db	978	TQ-----KILDRQFKNN-----DKTELEAK---L	1001
Qy	1174	EMNARSLOQKLETERELK-----ORLBEQA--KLOOQMDLOKNIHFLTGOLOEALDRA	1226
Db	1002	QSVANQSSASTLDRPSKNGNNGYUDDQTPREFHELDVGRG-----DKS	1049
Qy	1227	DLTTERSDELYOLENTIOVLYSHEKVKMEGTISQOTKLIDFLQAKMDQPAKKKVPLOYN	1286

Dh 1050 KKAFTKRRDSF-----FGLKDRLSGRGRSÖK-----RSKSDVYENN 1088
Qy 1287 ELKALKEKARCAELBEAL-----ÖKTRIELRSAREBAARAKATDHPHSTPATAR 1338
Dh 1089 QL-----LEBAVSLPPSRDPSTRYAVASSNDKYRETHSVGSGESTKSLYQ 1135
Qy 1339 QQLAMSAIVSPHEQPSAMSLAPPSRRKSSPPPEPSRLK--ERMHNHIIPIRFVGL 1396
Dh 1136 H---STLIELDODDKAKYFLIPPAALNEPA-----SLMRKCKDLHYNDHTF--VAV 1185
Qy 1397 NMR-ATCAVCLDVTN--FGRÖASKCLEQVWCHPKCSTCLPATCGLPAYATHF--TEAF 1452
Dh 1186 KVKGAGATCNVQÖQRIRSSFSÖQAVQCRDCKVCHKTC-----HYKIDAF 1229
Qy 1453 C 1453
Dh 1230 C 1230
RESULT 13
T16507
hypothetical protein F59A6.5 - Caenorhabditis elegans
C/Species: Caenorhabditis elegans
C/Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 09-Jul-2004
C/Accession: T16507
R,Nhan, M.
submitted to the EMBL Data Library, December 1995
A/Description: The sequence of C. elegans cosmid F59A6.
A/Reference number: Z18526
A/Accession: T16507
A/Status: preliminary; translated from GH/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-1128 <NHA>
A/Cross-references: UNIPROT:Q21025; EMBL:U41994; NID:G1123047; PID:G1123050; PIDN:AAA834
A/Genes: CESP:F59A6.5
A/Intons: 35/3; 335/3; 685/3; 973/3; 1097/3
Query Match 7.28; Score 751; DB 2; Length 1286;
Best Local Similarity 21.64; Pred. No. 1.9e-13;
Matches 299; Conservative 289; Mismatches 583; Indels 214; Gaps 50;
Qy 640 ELDEKLEKAVKASTAEATELLÖNTÖRÖKEREARELEKÖNREDSSEGRKKLVAEERRHS 699
Dh 42 ESHEDLKRRV---LDAENITÖDLKSRDALH---ETLVDAKGLAESV---TIEQSKVST 92
Qy 700 LENKVKLETFMERENRLKDDIQTKSQÖIQÖMADKILEBKREAOVSAQHLEVHLKÖK 759
Dh 93 QETRIYR-----RDVNLLEDLKHQSQIRILÖNKKCTLEMEKOTLÖETIÖRÖDDCKET 147
Qy 760 EÖHYE---EKIKVLDNÖ-----IKDÖLADK--ETLENMÖRHEEBAHEKGLISEQ 805
Dh 148 ETELESRSRLHYLEKELSAKANDIFPVYTKDLDHKNELTSPREYVTKLSEANR---EK 204
Qy 806 KAMINMDSKIRISLEQÖIVELSEANKLAANSULTÖNMMKÖAEIMSELÖQÖFYLE--T 863
Dh 205 KA-----LEEKLEKRYKNDMKENDRKSLELNÖYVTTÖVLSLEVÖLSAHEFEFT 253
Qy 864 QÖQKLEAÖNKLBE--ÖLE--KISHÖQSD--KNRLL--ELETRILREVSLEHEÖKLELK 915
Dh 254 PVRKMSKIRELDEYHÖLSAKVIESMNDLKIKNETITKELSDGTBLVKNKEE--LEDL 311
Qy 916 RÖLTELÖLS-----LÖERESÖTALÖA-----APALESÖLÖAKTELEETTA 959
Dh 312 RÖTTTASIGDSEÖATKYLHEENMKLTÖKADIRCELLAEARKVEG--FDKLEKÖLEKERD 370
Qy 960 ABEIEÖALTÖHÖRDEIQÖKPPALRNSCTVITDLBEÖQNLÖEDNAELNÖNÖFYLSKÖDEA 1019
Dh 371 ALADVÖKIRERVKKNVERELÖSL--TSLMAERÖÖLE-----ELKTKMFSEPMIKÖH 420
Qy 1020 SGANDEIVÖRSEVDHIREITEREMÖLTSÖKÖTMEALKT-----CTMLEÖÖVMDLE 1072
Dh 421 ESAKNELSRÖEÖKLDÖMÖKHLIMADÖQCSIFKSLKESABESRRARALÖCÖNEMVVRIDÖ 480

Qy 1073 ALNDELLEKÖRÖWEANRSVLDGKESÖFEGCYRELÖMÖLDEKÖSAPARADÖRITESQÖVE 1132
Dh 481 T-----SLESÖR-----KVEEVEMLKAEANSRÖAKKLEFMEKEBIEÖEVH 518
Qy 1133 LAVENKAEILALÖQALKEÖ-----KLKAESIDKLENDLEKKHMLÖNNARSILÖQÖKLET 1186
Dh 519 L---DYRÖELSRLAETKKGEDADHRLTISQÖDSSELRSKAKTIÖEVKADNÖQVÖMLVE 575
Qy 1187 ERELKÖRLLEBÖKÖLÖQÖDÖLÖKNHIFRITÖGLÖEALDRADLLKTERSÖLEYÖENIÖVL 1246
Dh 576 VRÖHÖEKILEENVRILRKGMADALAKIEEYKRSVÖNSÖETCERLERSATÖEDLIDLE-- 633
Qy 1247 YSHETKMEGTISÖQÖTKLIDFLÖAKMDÖPÖKKKVVÖLÖVNELEKALKEKARCAELBEAL 1306
Dh 634 --BELÖEKÖQÖLLESKELVYTHLSÖID--AKÖTÖP-----KÖRNSLSTLSTVE-- 679
Qy 1307 ÖKTRIELRSAREBA--HRKATDHPHSTPATARÖQLAMSAIVRSPEHÖPSAMSLAPP 1363
Dh 680 MÖTSVYMRBAEVRALBÖQÖRL---MENLAEKRÖ-----LVDSKKSÖSTANTTIVT 730
Qy 1364 SSRR--KESSTPEPSRLKERMHNIPIRFNV--GLNFRATCAVCLDTVHFGÖQÖSKC 1419
Dh 731 TTTEISKSQSÖSSELSNR--ÖGTREHDIPIHKWAFRHVGYLSMKCSLCFVGISAFÖAKÖKC 789
Qy 1420 LECÖVWCHPKCSTCLPATCGLPAYATHFTEAFRCRKNKNSPGLÖTEPESSLHLBGMÖKV 1479
Dh 790 SHÖDVAHVHASCARVNVNCTGMPÖQÖCAYVREN--HTTVSSSGV-----SEGRMNMÖLKV 841
Qy 1480 PRNKRQÖQÖQMDKYVLESGKVLIDYNEAREAGÖRPVEFEFTCLDPDÖVSIHGAÖGAS 1539
Dh 842 YRDÖMTGÖST--WASWAMMÖLTRISFTYNDGADL--EKRFPSID--NÖGÖVNLRTGÖMÖPD 898
Qy 1540 LANTAKADVPIYILKESHPHITTCWPGRTLYLLASFPDÖKÖRVÖTALÖSVAGÖVREKA 1599
Dh 899 CDSWÖANNVNLMKM-----PRCÖLYIILASQPSARRMÖECLÖTÖRKKMMLNSK 949
Qy 1600 EADAKLNSILKLEBDÖDLDMÖCTLPESDÖVVLVTEBGLVYANLÖK--NSLTHVÖGIGA 1658
Dh 950 SSIAEF--TCLLVNLSPPNMLKIFKAYTIDW--ILFATÖTGLÖFTSISÖPRNPRIGÖPNS 1006
Qy 1659 VFÖIYITIKLEKMLTAGEBALCLVDVKKVQÖSLQÖSHLPÖPÖDISPINIFÖAVKÖCHLP 1718
Dh 1007 VTSLEIMAEIKCYAMVANSÖQÖLAMIPL-----DSLILMÖSTÖPÖSRPEILPEFEHVNI 1062
Qy 1719 GAKIENGÖLÖICAMPSKVILLRYNENLSKYCIKEIETSEPCSHFTNYSILIGTNKE 1778
Dh 1063 KYHÖQÖGÖRFLLISDÖTHLHVRYNATRDIFÖQYAKFDVPEPISFIESAPSGIIFACDTE 1122
Qy 1779 YEIDMKÖYTLBEFLDÖNDISLAPAYAPASNSFPVSIÖVÖNSAGÖBEVILCTHERGÖRV 1838
Dh 1123 YVVALDÖHÖT-----SSNVASARKLMSPKNPEPISÖMÖIN---ÖNEVILAVYÖNGÖIFV 1171
Qy 1839 DSVGRSRTDÖKWSRLPAPAYREPLYFVTFHNSLEVEIÖIÖARSAGTAP--RAYLDIPN 1897
Dh 1172 NLHÖQSÖRNETTEMKMPHEFITTSFPLYIYHDDSIÖIÖIÖISTÖRYTALÖRÖALPECVN 1231
Qy 1898 PRYIGPÖLISGÖIYASÖYÖDKLRYVLCÖGNLVKESGTEHNR-----GÖTSRSSSPNKG 1952
Dh 1232 AHVIGRÖYGVÖLISVSN-----DSTEVHPSÖTATGÖKÖKNVSKRG 1274
Qy 1953 PPTYN 1957
Dh 1275 ASPEYN 1279
RESULT 14
T26101
hypothetical protein W02B8.2 - Caenorhabditis elegans
C/Species: Caenorhabditis elegans
C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C/Accession: T26101
R,Sime, M.

submitted to the EMBL Data Library, October 1996

A:Reference number: Z20151

A:Accession: T26101

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-1256 <WIL>

A:Cross-references: UNIPROT:Q9VNH4; EMBL:Z81136; P1DN:CAB03458.1; GSPDB:GN00020; CESP:W02B8

A:Experimental source: clone W02B8

C:Genetic:

A:Gene: CESP:W02B8.2

A:Map position: 2

A:Introns: 27/3; 327/3; 670/3; 949/3; 1073/3

Query Match 6.9%; Score 728; DB 2; Length 1256;

Best Local Similarity 20.6%; Pred. No. 7.5e-13;

Matches 307; Conservative 304; Mismatches 540; Indels 338; Gaps 54;

QY 545 QSRKLOEKEOEVOAQUEERKIMNOLEEDVVSARRSDIYESRLAAEFKKAT 604

DB 10 QSKTPKSKARISSPYQVIDL--NNESQEDKTRTLAENIIODLKSERDLHE----- 62

QY 605 ECHKLLKAKDQKPEVEYAKLEKINAEOQLKIQLOEKLEKAVKASTATELLQN--- 661

DB 63 -----SLVDKAGLNSVII-----ERENKYSTQETRIYRDDVT 95

QY 662 -TQAKERARELEKLN-----EDSEGIKKLVAEERHSLNKKVLET 709

DB 96 LLEBDLQKQESQRIILNRCILRETEKQKQMDITISGYQEDLKENE-----IRIEN 145

QY 710 MERRRNLKODIQTKSQI-----QOMADKILLEEKRE-----AOVSAGHLEVHLKQKE 760

DB 146 LNSRLHLEBELAKTHEIPISEBELNKTWKLNKNSQFQTKLAETISSNR--NLERKV 203

QY 761 OHYEKKIKVLDNOIKKDLADKETLNNMORHEEBAHEKGLISEOKAMINAM--SKIRS 818

DB 204 QKREBELIVD-----QSLFVHQDQENQ--KVLKENVQQLSDRLDYLTPKRX 249

QY 819 LEORIYELSEANKLANSSILFTQPNMKAQEMISLSEKQKYLETQAGKLEAONRKLEQ 878

DB 250 DVSRKERD-----FLQPSAKIIEETSELKKNARLERISE-----KEE 291

QY 879 LEKISHQDSKRNLELETRLEVEISLEHE-----QCLEKRLQUTELQSLQERESQ 931

DB 292 LVKVT-----KELELOKTVTQANGDSQATKYLHAENMKLTRLQADLCLEERRN 345

QY 932 LFLALQARALAESQLQAKTELEETT--AAEEEIOALTAHDEIORKFDALRNSC--- 985

DB 346 LKGFDEKRELEKQRBDALEDVRIETELKKNVEIELRSLKILABERBEQIDELKRYAGY 405

QY 986 -TVITDLEBQNLQITEDNLANNNQNFYLSKQDDEASGANDIYOLNSEVHLRREITERE 1044

DB 406 EVLRHDEAVKNELEAKBEKLNMGAL-----VMADQSSHFK---TLKE 448

QY 1045 MQLTSQKQTEALAKTCTMLEBOVMDLEANDELLEKROWEMRSLVGDQKQFQRCVR 1104

DB 449 TAEGRSRRAL-----QCNEMVARIRGLE-----SLEQKQVTE-----Q 483

QY 1105 ELQRMIDTEKQSPARADORTESROYVELAVKEHKAIIALLQALKEKQKALASLSD-KL 1163

DB 484 ELE-WKAENVRQAKKIEPKKEIQLTHLDYRELSKL-----AKGGSGHEASQORSEL 537

QY 1164 NDLEKHAMLENNARSLOQLETERELKORLEEOAKLQOQOMDLOKNNHIFRLTQGLQEA 1223

DB 538 RSKAKTTIOEVKADKKNKQQLLEVRNQSKVLEENYKLR-----KGMALAI 583

QY 1224 DRADLKT-----ERSDLEYOLENIOVLYSHEKVMKMG--TISQQTQLDLPLOAKMD 1273

DB 584 EKLEEFKRWMSHRSREAGERLOLEAKNEBEKVLKVEBELQEKRELEVKEKEELVNYLOSQIN 643

QY 1274 -QPAKKKVPIQVNELEKLEKAR-CALEE--ALOKRIRLEARSAREAAHKKYTDHP 1329

DB 644 TKOTKQPKLSRSTLMSTISSEVDTSTYVREVEVRALREOREELQAYLAE--KKKADLO 701

QY 1330 HPSPTARQOIMASATVSPHQPSAMSLAPSSSRKSSSTPEERSRLKERMHNIP 1389

DB 702 KSRSTANTTLLTIST--ATEVSKACELSTYPAT-----MRDIP 740

QY 1390 HRFVGLNMR-----ATKCAVCLDITVFGQASKLECOVMCHPKSCSTCLPATGCLPAEY 1444

DB 741 HKM-----LEFHHYSLSKKSLCFVGITFFPKANKCSICDVRAHASCAPRVANNCTGIPAC 797

QY 1445 ATHTFAFCRDKNNSPGLQTEPSSSLHLEGMKVPNNNRGQGWDRKYIVLEGSRYLI 1504

DB 798 ATYYQDA---QAVGGIG-----RMNGWLRVYRDD--LPSWTWLSLWAMDSNHYKF 843

QY 1505 YDNBARAGRPVEBEL-----CLPDG--DVSFHGAVGASELANAKADVPYILKMSHP 1558

DB 844 YRDAGADNLENPFITIDLNKEQWILRTGQVVALPGDMRVLVT-----IKLOT-- 891

QY 1559 HTTCWPGRTLYLAPFPDPKQWVTALLESVAVAGRVSRKAEADAKLGLSLKLEGGDR 1618

DB 892 -----RSVHYIAPFPKSAERMACTQNAQTRMKNSKPSIAEY--SCLTISLPNN 942

QY 1619 LDMNCTLPFSQVVLVGTBEGYALNVLK--NSLTHVPQIGAVFOIYIINDLEKLMIAE 1677

DB 943 LKIFKAHTIEDW--ILFATQGLFETISQPRNPTRIAGPSVTSLEVMSEINCVMATNS 1001

QY 1678 ERALCLVDYKVKQSLAQSHLPAPQDISPVIFPAVKGCLFGAKTENGCTCAAPSKV 1737

DB 1002 NROLALIPLDLSLTAQOSTH---PSIRAEVLVEFGVHTIRYHQDQGRFLVSDPTOL 1057

QY 1738 VILRVNENLSKYCIRKEIETSEPCSCHEFTNYSILGTNKFYELDMKQYILLEFLDNDH 1797

DB 1058 HIRKINSTRDVFHAFKLVVPEVPSFTESTPACFIRASDFYVPL-----DH 1105

QY 1798 SLAPVAFAS-----SNSPFSIVQVNSAGQREYLLCFEFQGVFVDSYGRRSRTDLDK 1852

DB 1106 QSPSDVSARKLMPRRSDVPVSAHALTA---NEILLAYONHIFVNLVYEGSRNQIIEW 1161

QY 1853 SRULPAPAREPYLFTVTHNSLEVEIEIQARSSAGTPARAYLIDIPNRYLPAISGAYL 1912

DB 1162 EKMPIEFAYTSPLLTYIYHDSIIEVOISK-----SSKETVL 1197

QY 1913 ASSYODKLVYICCKNLVKESSG-----TEHHRGPTSSSPNKR 1951

DB 1198 A---EREVLACVANAHIQSDGVLSVSKDITEVHRFSATCGTGYKR 1242

RESULT 15

G86431 protein kinase T518.9 protein - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004

C:Accession: G86431

R:Theologas, A.; Ecker, J. R.; Palm, C. J.; Federspiel, N. A.; Kaul, S.; White, O.; Alonso, Chien, C. W.; Chung, M. K.; Conn, L.; Conway, A. B.; Conway, A. R.; Creasy, T. H.; Dewar, K.;

ansen, N. E.; Hughes, B.; Hultzer, L.

Nature 408, 816-820, 2000

A:Authors: Hunter, J. L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.;

Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

C.A.; Li, J. H.; Li, Y.; Liu, S. X.; Liu, Z. A.; Lueros, J. S.; Maltf, R.; Marziani, A.;

Authors: Salzberg, S. L.; Schwartz, J. R.; Shim, P.; Southwick, A. M.; Sun, H.; Tallon, I.

Ker, M.; Wu, D.; Yu, G.; Frazer, C. M.; Venter, J. C.; Davis, R. W.

A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A:Reference number: A86141; MUID:21016719; PMID:11130712

C:Accession: G86431

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-522 <STO>

A:Cross-references: UNIPROT:Q9SA79; GB:AE005172; NID:94587520; P1DN:AAD25751.1; GSPDB:GN

C:Genetic:

A:Map position: 1

C:superfamily: Arabidopsis thaliana protein kinase T518.9; protein kinase homology

Query Match 6.9%; Score 721; DB 2; Length 522;

Best Local Similarity 35.9%; Pred. No. 5.2e-13;

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 11, 2005, 16:44:55 ; Search time 204 Seconds

(without alignments)
3363.442 Million cell updates/sec

Title: US-10-791-666-2

Perfect score: 10493

Sequence: 1 MLKFKYGARNPLDGAAPL.....AGAVRPLSGVKNKWDQSSV 2054

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1432185 seqs, 334051727 residues

Total number of hits satisfying chosen parameters: 1432185

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
- 6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
- 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
- 9: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep:*
- 10: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*
- 11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*
- 12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
- 13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep:*
- 14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*
- 15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*
- 16: /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pep:*
- 17: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
- 18: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep:*
- 19: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
- 20: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	10493	100.0	2054	US-10-028-946-2	Sequence 2, Appl1
2	10486	99.9	2054	US-10-415-011-21	Sequence 21, Appl1
3	10464.5	99.7	2055	US-10-618-941-67	Sequence 67, Appl1
4	10458.5	99.7	2053	US-09-964-956-11	Sequence 11, Appl1
5	10458.5	99.7	2053	US-10-262-511-2	Sequence 2, Appl1
6	10425.5	99.4	2066	US-09-964-956-9	Sequence 9, Appl1
7	10425.5	99.4	2066	US-10-262-511-14	Sequence 14, Appl1
8	10109.5	96.3	2055	US-10-017-216-4	Sequence 4, Appl1
9	10022.5	95.5	2053	US-10-017-216-2	Sequence 2, Appl1
10	10022.5	95.5	2053	US-10-325-430-12	Sequence 12, Appl1
11	10022.5	95.5	2053	US-10-757-262-52	Sequence 52, Appl1
12	9916	94.5	1958	US-10-028-946-4	Sequence 4, Appl1
13	8161	77.8	1641	US-09-964-956-40	Sequence 40, Appl1

14	8161	77.8	1641	13	US-10-017-216-5	Sequence 5, Appl1
15	7935	75.6	1597	11	US-09-964-956-41	Sequence 41, Appl1
16	7935	75.6	1597	11	US-10-017-216-6	Sequence 6, Appl1
17	7189	68.5	1441	15	US-10-412-897-3	Sequence 3, Appl1
18	6575	62.7	1286	11	US-09-964-956-38	Sequence 38, Appl1
19	6575	62.7	1286	13	US-10-017-216-7	Sequence 7, Appl1
20	5045	48.1	999	15	US-10-276-774-1487	Sequence 1487, Ap
21	4871	46.4	940	11	US-09-964-956-39	Sequence 39, Appl1
22	4601	43.8	883	11	US-09-964-956-37	Sequence 37, Appl1
23	4115	39.2	832	15	US-10-262-511-6	Sequence 6, Appl1
24	4097.5	39.0	847	15	US-10-262-511-8	Sequence 8, Appl1
25	3244	30.9	623	15	US-10-262-511-10	Sequence 10, Appl1
26	3226.5	30.7	638	15	US-10-262-511-12	Sequence 12, Appl1
27	3165	30.2	623	15	US-10-262-511-4	Sequence 4, Appl1
28	2440	23.3	497	14	US-09-804-471A-2	Sequence 2, Appl1
29	2440	23.3	497	14	US-10-238-709-2	Sequence 2, Appl1
30	2440	23.3	497	15	US-10-724-594-2	Sequence 7, Appl1
31	2435	23.2	497	15	US-10-311-034-7	Sequence 4, Appl1
32	2175.5	20.7	494	9	US-09-804-471A-4	Sequence 4, Appl1
33	2175.5	20.7	494	14	US-10-238-709-4	Sequence 4, Appl1
34	2175.5	20.7	494	15	US-10-724-594-4	Sequence 4, Appl1
35	1792	17.1	373	15	US-10-120-835-42	Sequence 42, Appl1
36	1655	15.8	319	15	US-10-412-897-2	Sequence 2, Appl1
37	1509.5	14.4	1770	15	US-10-433-794-19	Sequence 19, Appl1
38	1489	14.2	1719	15	US-10-288-798-2	Sequence 2, Appl1
39	1489	14.2	1719	15	US-10-362-892-2	Sequence 2, Appl1
40	1487.5	14.2	1732	17	US-10-679-366-2	Sequence 2, Appl1
41	1467.5	14.0	1738	15	US-10-210-130-100	Sequence 100, App
42	1466	14.0	1664	15	US-10-210-130-102	Sequence 102, App
43	1439	13.7	1711	17	US-10-781-581-206	Sequence 206, App
44	1438	13.7	1711	9	US-09-771-161A-219	Sequence 219, App
45	1438	13.7	1711	9	US-09-771-161A-220	Sequence 220, App

ALIGNMENTS

RESULT 1
US-10-028-946-2
; Sequence 2, Application US/10028946
; Publication No. US20020123622A1
GENERAL INFORMATION:
APPLICANT: Yu, Xuanchuan
APPLICANT: Miranda, Maricar
TITLE OF INVENTION: No. US20020123622A1 Human Kinases and Polynucleotides Encoding
FILE REFERENCE: LEX-0289-USA
CURRENT APPLICATION NUMBER: US/10/028,946
PRIOR FILING DATE: 2001-12-20
PRIOR APPLICATION NUMBER: US 60/258,335
NUMBER OF SEQ ID NOS: 4
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 2
LENGTH: 2054
TYPE: PRT
ORGANISM: homo sapiens
US-10-028-946-2
Query Match 100.0%; Score 10493; DB 13; Length 2054;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2054; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MLKFKYGARNPLDGAAPL...ASR...LTF...G...P...T...Q...M...S...L...R...E...G...I...D...L...P...V...L...F...E...60
Db 1 MLKFKYGARNPLDGAAPL...ASR...LTF...G...P...T...Q...M...S...L...R...E...G...I...D...L...P...V...L...F...E...60
QY 61 ECGSOPALMKIKHVSNFVRKYSDTLAELOELQPSAKDFEVSILVCGHFAEVQVVRERATG 120
Db 61 ECGSOPALMKIKHVSNFVRKYSDTLAELOELQPSAKDFEVSILVCGHFAEVQVVRERATG 120
QY 121 DTYAMKMKKALLAQGVSPFEEERNITLSRSYSPWLPOLQVAFODKHLVLMVEYPPG 180
Db 121 DTYAMKMKKALLAQGVSPFEEERNITLSRSYSPWLPOLQVAFODKHLVLMVEYPPG 180

Db 121 DIYAMVKKKALLAQEVSFEEERNILSSTSPMIPOLQYAFODKNHLYLVNEVOPGC 180
 Qy 181 DLASLNRVEDOLDENILQFYLAELILAVSHLVHMGVYHRDIKPBENILVDRTHGHIKLVDF 240
 Db 181 DLASLNRVEDOLDENILQFYLAELILAVSHLVHMGVYHRDIKPBENILVDRTHGHIKLVDF 240
 Qy 241 GSAAKNSNNKVNAKLPIGTPDYMAPEVLYVMNGDGKTYGLDCDWMVGVYIAYEMTYGR 300
 Db 241 GSAAKNSNNKVNAKLPIGTPDYMAPEVLYVMNGDGKTYGLDCDWMVGVYIAYEMTYGR 300
 Qy 301 SPFAEGSARTENNIMNFORLKFPPDPKVSDFLDIQLSLCGOKERLKEGCLCPHFF 360
 Db 301 SPFAEGSARTENNIMNFORLKFPPDPKVSDFLDIQLSLCGOKERLKEGCLCPHFF 360
 Qy 361 SKIDMNNIRNSPPFPVPTLKSDDDTSNFDEPEKNSWVSSPCQSPSGSGEBELPFVGS 420
 Db 361 SKIDMNNIRNSPPFPVPTLKSDDDTSNFDEPEKNSWVSSPCQSPSGSGEBELPFVGS 420
 Qy 421 YSKALGILGRSESVSGSDSPAKTSSMEKULLIKSKELQDSQDYCKHKEQMTLHRRVS 480
 Db 421 YSKALGILGRSESVSGSDSPAKTSSMEKULLIKSKELQDSQDYCKHKEQMTLHRRVS 480
 Qy 481 EYEAVLSQKEVELKASETORSLBODLATYTCESSLKRSLEQARMESVQEDDQALQULH 540
 Db 481 EYEAVLSQKEVELKASETORSLBODLATYTCESSLKRSLEQARMESVQEDDQALQULH 540
 Qy 541 DIRQSRKLOEIKQEOYQAOVEENRLMMNQLEBDLVASARRSDYSESLERSLAAEFK 600
 Db 541 DIRQSRKLOEIKQEOYQAOVEENRLMMNQLEBDLVASARRSDYSESLERSLAAEFK 600
 Qy 601 RKATECOHKLKAKDOQGEVGEYAKLEKINAEOOLKIQELQEKLEKAVKASTATELILQ 660
 Db 601 RKATECOHKLKAKDOQGEVGEYAKLEKINAEOOLKIQELQEKLEKAVKASTATELILQ 660
 Qy 661 NIRAARARERELKQNRDSSSGIRKUYEABERHSLENKYRLJEMERRRNRLLKDD 720
 Db 661 NIRAARARERELKQNRDSSSGIRKUYEABERHSLENKYRLJEMERRRNRLLKDD 720
 Qy 721 IQTSQOIQOMADKILLEEGHREAOVSAQHLFVHLKQKEQYBEKIKVLDNQIKCLAD 780
 Db 721 IQTSQOIQOMADKILLEEGHREAOVSAQHLFVHLKQKEQYBEKIKVLDNQIKCLAD 780
 Qy 781 KETLENNMQRHEEBAHEKGLISLQKAMINAMDSKINSLEORIVELSEANKLAANSILFT 840
 Db 781 KETLENNMQRHEEBAHEKGLISLQKAMINAMDSKINSLEORIVELSEANKLAANSILFT 840
 Qy 841 ORNNKAQEBMISELRQOQFYLETQAGKLEAQRKLEQLEKISHODSHDKRRLLETRL 900
 Db 841 ORNNKAQEBMISELRQOQFYLETQAGKLEAQRKLEQLEKISHODSHDKRRLLETRL 900
 Qy 901 REVSLEHEBQKLEIKROLTELQSLQERESQTLAQARALAESQLQAKTELEETTAEA 960
 Db 901 REVSLEHEBQKLEIKROLTELQSLQERESQTLAQARALAESQLQAKTELEETTAEA 960
 Qy 961 EEBIOLATLHARDEIQRKFDALRNSCTVITDLEEOQLNQTEDNABINNPFYLSKQDBAS 1020
 Db 961 EEBIOLATLHARDEIQRKFDALRNSCTVITDLEEOQLNQTEDNABINNPFYLSKQDBAS 1020
 Qy 1021 GANDEIYQLRSEVHLRREITEREMQULSOKQWEMALKTTCTMEEBQWMDLEANDLELE 1080
 Db 1021 GANDEIYQLRSEVHLRREITEREMQULSOKQWEMALKTTCTMEEBQWMDLEANDLELE 1080
 Qy 1081 KEROMEAMRSVLAGBEKSQFECRVELORMLDTEKOSRARADORTTESQOVELAVKEKA 1140
 Db 1081 KEROMEAMRSVLAGBEKSQFECRVELORMLDTEKOSRARADORTTESQOVELAVKEKA 1140
 Qy 1141 EILALQALKEQKUKAESLSDKLNLEKQAMLENNASLQOKLETERELKQRLLEBQAK 1200
 Db 1141 EILALQALKEQKUKAESLSDKLNLEKQAMLENNASLQOKLETERELKQRLLEBQAK 1200
 Qy 1201 LOQOMDLOKNHIFLUTQGLQALDRADILKTERSDULEYQLENIQVLYSHEKVKKEGITSQ 1260
 Db 1201 LOQOMDLOKNHIFLUTQGLQALDRADILKTERSDULEYQLENIQVLYSHEKVKKEGITSQ 1260

Qy 1261 QTKLIDFLQAKMDQPAKKKVPLOYNELKLALKEKARCALEBALQKTHIELRSAREEA 1320
 Db 1261 QTKLIDFLQAKMDQPAKKKVPLOYNELKLALKEKARCALEBALQKTHIELRSAREEA 1320
 Qy 1321 AHRKATDHPHSTPATRQOILAMSAIVRSPEHOPSAMSLAPSSRRKESSTPEEFSRL 1380
 Db 1321 AHRKATDHPHSTPATRQOILAMSAIVRSPEHOPSAMSLAPSSRRKESSTPEEFSRL 1380
 Qy 1381 KERHNNHI PHRFVNGLMRATKCAVCLDTHVFGROASKLECOVMCHPKSTCLPATGCL 1440
 Db 1381 KERHNNHI PHRFVNGLMRATKCAVCLDTHVFGROASKLECOVMCHPKSTCLPATGCL 1440
 Qy 1441 PABYATHPTAFGRDKNSNGLOTKEPSSSLHLEGMKVKVRNNKRGQGGMDRKYIVLEGS 1500
 Db 1441 PABYATHPTAFGRDKNSNGLOTKEPSSSLHLEGMKVKVRNNKRGQGGMDRKYIVLEGS 1500
 Qy 1501 KVLVYDNEAREAGORPVEEFELCLPGDVSIGHAVGASELANAKADVPIYILKMHESHPT 1560
 Db 1501 KVLVYDNEAREAGORPVEEFELCLPGDVSIGHAVGASELANAKADVPIYILKMHESHPT 1560
 Qy 1561 TCMRGRTLYILAPSPDKQRMVTLAESVAVGHRVSRKADAKLIGNSILKLEGGDRLD 1620
 Db 1561 TCMRGRTLYILAPSPDKQRMVTLAESVAVGHRVSRKADAKLIGNSILKLEGGDRLD 1620
 Qy 1621 MCTLPSPDQVVLVGTBEGLYALNVLKNSLTHVPGICAVFOIYI IKDLEKLMIAGERA 1680
 Db 1621 MCTLPSPDQVVLVGTBEGLYALNVLKNSLTHVPGICAVFOIYI IKDLEKLMIAGERA 1680
 Qy 1681 LCIVDYKVKVQSLAQSHLPAOPDISPVIPEAVGCHLFGAGKLENGCLCAAMPKVVIL 1740
 Db 1681 LCIVDYKVKVQSLAQSHLPAOPDISPVIPEAVGCHLFGAGKLENGCLCAAMPKVVIL 1740
 Qy 1741 RYENENLSKYCIKREIETSEPCSCIHFTNYSILIGTNKFYEIDMKOYTLLEBFLDKNDHSLA 1800
 Db 1741 RYENENLSKYCIKREIETSEPCSCIHFTNYSILIGTNKFYEIDMKOYTLLEBFLDKNDHSLA 1800
 Qy 1801 PAVPAASNSFPVSIQVONSAGOREEYLLCFHEBGVFDVDSYGRSRDIDLKMSRLPAFA 1860
 Db 1801 PAVPAASNSFPVSIQVONSAGOREEYLLCFHEBGVFDVDSYGRSRDIDLKMSRLPAFA 1860
 Qy 1861 YREBYLPTHFNLSLEVEIQARSSAGTPARAYLIDIPRVLGPALISSGAYLASSYODKL 1920
 Db 1861 YREBYLPTHFNLSLEVEIQARSSAGTPARAYLIDIPRVLGPALISSGAYLASSYODKL 1920
 Qy 1921 RVICCKGNLVKESGTEHHRGPSTSRSSPNKRGPPTYNEHITTKVASSPAPBEGSHPREP 1980
 Db 1921 RVICCKGNLVKESGTEHHRGPSTSRSSPNKRGPPTYNEHITTKVASSPAPBEGSHPREP 1980
 Qy 1981 STPHRYEGRTLELRDCKSPERPLEREKSPGRIILSTRERNSPAPLFEDESSRGRLPAGAVRT 2040
 Db 1981 STPHRYEGRTLELRDCKSPERPLEREKSPGRIILSTRERNSPAPLFEDESSRGRLPAGAVRT 2040
 Qy 2041 PLSQVNVKVMQSSV 2054
 Db 2041 PLSQVNVKVMQSSV 2054

RESULT 2
 US-10-415-011-21
 ; Sequence 21, Application US/10415011
 ; Publication No. US20040053394A1
 ; GENERAL INFORMATION:
 ; APPLICANT: INCYTE CORPORATION
 ; APPLICANT: GURURAJAN, Rajagopal
 ; APPLICANT: BAUGHEY, Mariah R.
 ; APPLICANT: CHAWLA, Nandinder K.
 ; APPLICANT: ELIOTT, Vicki S.
 ; APPLICANT: XU, Yuming
 ; APPLICANT: ARVIZU, Chandra S.
 ; APPLICANT: YAO, Montique G.
 ; APPLICANT: RANKOWAR, Jayalaxmi
 ; APPLICANT: DING, Li

APPLICANT: TANG, Y. Tom
APPLICANT: HAPALA, April J.A.
APPLICANT: NGUYEN, Daniel B.
APPLICANT: GANDHI, Ameena R.
APPLICANT: LU, Yan
APPLICANT: YUE, Henry
APPLICANT: BURFORD, Neil
APPLICANT: BANDMAN, Olga
APPLICANT: TRIBOULEY, Catherine M.
APPLICANT: LAL, Preeti G.
APPLICANT: RECIPON, Shirley A.
APPLICANT: LU, Dyung Aina M.
APPLICANT: BOROMSKY, Mark L.
APPLICANT: THORNTON, Michael B.
APPLICANT: SWANAKER, Anita
APPLICANT: THANGAVELOU, Kavitha
APPLICANT: KHAN, Farrah A.
APPLICANT: ISON, Craig H.
TITLE OF INVENTION: HUMAN KINASES
FILE REFERENCE: PI-0262 USN
CURRENT APPLICATION NUMBER: US/10/415, 011
CURRENT FILING DATE: 2003-04-18
PRIOR APPLICATION NUMBER: PCT/US01/47728
PRIOR FILING DATE: 2001-10-20
PRIOR APPLICATION NUMBER: US 60/242,410
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: US 60/244,068
PRIOR FILING DATE: 2000-10-27
PRIOR APPLICATION NUMBER: US 60/245,708
PRIOR FILING DATE: 2000-11-03
PRIOR APPLICATION NUMBER: US 60/247,672
PRIOR FILING DATE: 2000-11-09
PRIOR APPLICATION NUMBER: US 60/249,565
PRIOR FILING DATE: 2000-11-16
PRIOR APPLICATION NUMBER: US 60/252,730
PRIOR FILING DATE: 2000-11-22
PRIOR APPLICATION NUMBER: US 60/250,807
PRIOR FILING DATE: 2000-12-01
NUMBER OF SEQ ID NOS: 44
SOFTWARE: PERL Program
SEQ ID NO 21
LENGTH: 2054
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
OTHER INFORMATION: Incyte ID No. US20040053394A1 7484498CD1
US-10-415-011-21

Query Match 99.9%; Score 10486; DB 15; Length 2054;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2052; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

1 MKFKYARNPDLGAAEPPLASRSRLNLPFGKRPPTQOONSPLSREGILDALFLTFE 60
1 MLKFKYGARPLDGAEPPLASRSRLNLPFGKRPPTQOONSPLSREGILDALFLTFE 60
61 ECOSOPALMKIKHVSNFYKTSPTLAEOLOPSAKDFEVSILVCGHFAEYQVREKATG 120
61 ECOSOPALMKIKHVSNFYKTSPTLAEOLOPSAKDFEVSILVCGHFAEYQVREKATG 120
121 DIYAMKVMKKKALLAQOVSPFEERNTLSRSTSPMIPOLQYAFQDKNHLVLMVEYQPG 180
121 DIYAMKVMKKKALLAQOVSPFEERNTLSRSTSPMIPOLQYAFQDKNHLVLMVEYQPG 180
181 DLISLNRVEDQDENTLQYTLAELIIVHVSVHLMGVHRDIKENTILVDRGTGHIKLVDF 240
181 DLISLNRVEDQDENTLQYTLAELIIVHVSVHLMGVHRDIKENTILVDRGTGHIKLVDF 240
241 GSAAKONSNNKVVNAKLPIDTPDYMAPEVLTYMNGDGKGTGJDDCMKSVGVIAEMTYGR 300
241 GSAAKONSNNKVVNAKLPIDTPDYMAPEVLTYMNGDGKGTGJDDCMKSVGVIAEMTYGR 300

301 SPFAEGTSATFNNNMFORFLKFPDDPKVSSDLDIQSLCCQKRLKEGICCHPFF 360
301 SPFAEGTSATFNNNMFORFLKFPDDPKVSSDLDIQSLCCQKRLKEGICCHPFF 360
361 SKIDMNNIRNSPPFVFTLKSDDTSNPDPEKNSWSSSPCOLSPSGSGEELPVPVFS 420
361 SKIDMNNIRNSPPFVFTLKSDDTSNPDPEKNSWSSSPCOLSPSGSGEELPVPVFS 420
421 YSKALGILGRSESVSGIDSPAKTSMEKLLIKSKELQDSQDCHXMEQEMTLHRVS 480
421 YSKALGILGRSESVSGIDSPAKTSMEKLLIKSKELQDSQDCHXMEQEMTLHRVS 480
481 EYEAVLSQKEVELKASTQSLFQDLATYTESSTLKRSLQARMESVQEDDVALOLH 540
481 EYEAVLSQKEVELKASTQSLFQDLATYTESSTLKRSLQARMESVQEDDVALOLH 540
541 DIREQSRKLOEIKQOEYQAOVEEMRLMMNOLLEEDLVARRSDLYESLRSRLAAEFK 600
541 DIREQSRKLOEIKQOEYQAOVEEMRLMMNOLLEEDLVARRSDLYESLRSRLAAEFK 600
601 RKATECOHKLKAKDQCKPEVEYAKLEKINAEOQLIOELOELKRAVKASTATEBELQ 660
601 RKATECOHKLKAKDQCKPEVEYAKLEKINAEOQLIOELOELKRAVKASTATEBELQ 660
661 NIROAKERARELEKLNREDSSEGIKKLVEAERHSLNKKYKRLTERRRRLKDD 720
661 NIROAKERARELEKLNREDSSEGIKKLVEAERHSLNKKYKRLTERRRRLKDD 720
721 IQTSQOIQOMADKILTELEKREAOVSAGHLEVHLKQKQHEEKKVLDNQKKDLAD 780
721 IQTSQOIQOMADKILTELEKREAOVSAGHLEVHLKQKQHEEKKVLDNQKKDLAD 780
781 KETLENNMQHBEAEHKKILSEQKAMINAMSKIRLSQRIYELSEANKLAANSLSFT 840
781 KETLENNMQHBEAEHKKILSEQKAMINAMSKIRLSQRIYELSEANKLAANSLSFT 840
841 QRMNKAQEEWISLROQKFLYLETQAGKLEAQNRLKEQLKISHQSDKRNLELETRL 900
841 QRMNKAQEEWISLROQKFLYLETQAGKLEAQNRLKEQLKISHQSDKRNLELETRL 900
901 REVSLHEBEOKLELKLQTLTLOSLQRESQTLNQAARALBSQLAQTLELETTAA 960
901 REVSLHEBEOKLELKLQTLTLOSLQRESQTLNQAARALBSQLAQTLELETTAA 960
961 EEEIOALTARDEIORFDMRNSCTYITDLEEQNLQTDNLENNQNFYLSQQLDEAS 1020
961 EEEIOALTARDEIORFDMRNSCTYITDLEEQNLQTDNLENNQNFYLSQQLDEAS 1020
1021 GANDEIVQLRSEVDHLRREITEREMQLTSGQOTMBALKTTCTMLEQVMDLEALNDELLE 1080
1021 GANDEIVQLRSEVDHLRREITEREMQLTSGQOTMBALKTTCTMLEQVMDLEALNDELLE 1080
1081 KERQWEAMRSLVLDKESQFECRVRLEORMLDTEKQSAARDORTITERQVVELAVKHKA 1140
1081 KERQWEAMRSLVLDKESQFECRVRLEORMLDTEKQSAARDORTITERQVVELAVKHKA 1140
1141 EIALAOLAKQKQKAEESLSDKLNDLEKKHAMLEMMARSLQQLETREELKORLLEQAK 1200
1141 EIALAOLAKQKQKAEESLSDKLNDLEKKHAMLEMMARSLQQLETREELKORLLEQAK 1200
1201 LQOQMDLQKNHIFRLTQLOEALDRADLKTERTSDLEYOLENIQVLSHEKVMEGTISQ 1260
1201 LQOQMDLQKNHIFRLTQLOEALDRADLKTERTSDLEYOLENIQVLSHEKVMEGTISQ 1260
1261 QTKLIDFLQAKMDQPAKKKVPLOYNELKALKEKARCALEBEALQKTRIEKRSAREEA 1320
1261 QTKLIDFLQAKMDQPAKKKVPLOYNELKALKEKARCALEBEALQKTRIEKRSAREEA 1320
1321 AHRKATDHPHSTPATRQOILASATVRSPEHQPSAMSLAPSSRKESSTPEEFSRRL 1380
1321 AHRKATDHPHSTPATRQOILASATVRSPEHQPSAMSLAPSSRKESSTPEEFSRRL 1380
1381 KERWHEHNTPHRFNVGLNMRATKCAVCLDTVHFGQASKLEQCMGHFKCSTGLPATCGL 1440
1381 KERWHEHNTPHRFNVGLNMRATKCAVCLDTVHFGQASKLEQCMGHFKCSTGLPATCGL 1440

Db 1381 KERNHNIPIRHFVNGLMNRAIKCAVCLDVIHFGQASKLECYWCHPCKSTCULPATCGL 1440
Qy 1441 PAEVAHTEFAFCBCKMNSPGLQKBPSSSLHLEGMMKVPRNNKRGGQGMGRKTYIVEGS 1500
Db 1441 PAEVAHTEFAFCBCKMNSPGLQKBPSSSLHLEGMMKVPRNNKRGGQGMGRKTYIVEGS 1500
Qy 1501 KVLIDNEAREAGORPVYEFELCPDGDVSIHGA VGASELANAKADVPIYLKMEHPHT 1560
Db 1501 KVLIDNEAREAGORPVYEFELCPDGDVSIHGA VGASELANAKADVPIYLKMEHPHT 1560
Qy 1561 TCMFGRLLYLLAEPFPDQKRWYTLSESVAGGRVSRKAEADACLGNSLKLEGGDRLD 1620
Db 1561 TCMFGRLLYLLAEPFPDQKRWYTLSESVAGGRVSRKAEADACLGNSLKLEGGDRLD 1620
Qy 1621 MNCULPSPDOVVLVTEBGLVYALNVLKNSLTHVIGIGAFQIYIYIKOLEKLLMAGEERA 1680
Db 1621 MNCULPSPDOVVLVTEBGLVYALNVLKNSLTHVIGIGAFQIYIYIKOLEKLLMAGEERA 1680
Qy 1681 LCLVDVKKVQSLAQSHLPAOPDISPNI FEA VKGCHLFGAGKIBNGLCICAMPSTKVIL 1740
Db 1681 LCLVDVKKVQSLAQSHLPAOPDISPNI FEA VKGCHLFGAGKIBNGLCICAMPSTKVIL 1740
Qy 1741 RYNNENSKYCIKREIETSEPCSCIHFTNYSTLIGTNKRYEIDMKQYTLBEFLDKNDHSLA 1800
Db 1741 RYNNENSKYCIKREIETSEPCSCIHFTNYSTLIGTNKRYEIDMKQYTLBEFLDKNDHSLA 1800
Qy 1801 PAVFAASNSPFSVIVOVNSAGOREEYLLCFHEFGVFDVSGRSRTDCLKMSRLPLAFA 1860
Db 1801 PAVFAASNSPFSVIVOVNSAGOREEYLLCFHEFGVFDVSGRSRTDCLKMSRLPLAFA 1860
Qy 1861 YREBYLFEVTHENSLEVIEIQARSSAGTPARAYLDIPNRYLGPAISSGATYLAASYDKL 1920
Db 1861 YREBYLFEVTHENSLEVIEIQARSSAGTPARAYLDIPNRYLGPAISSGATYLAASYDKL 1920
Qy 1921 RVICCKGNLVKESGTEHHRGPSTSRSSPNKRGPTVNEHTKRVASSPAPPEGSHPREP 1980
Db 1921 RVICCKGNLVKESGTEHHRGPSTSRSSPNKRGPTVNEHTKRVASSPAPPEGSHPREP 1980
Qy 1981 STPHRYEGRTELPRDPSPGPLEREKSPGRILSTRERSPARLPEDSSRGRLPAGAVRT 2040
Db 1981 STPHRYEGRTELPRDPSPGPLEREKSPGRILSTRERSPARLPEDSSRGRLPAGAVRT 2040
Qy 2041 PLSQVNVKWDQSSV 2054
Db 2041 PLSQVNVKWDQSSV 2054

RESULT 3
US-10-618-941-67
; Sequence 67, Application US/10618941
; Publication No. US20040197792A1
; GENERAL INFORMATION:
; APPLICANT: WHYTE, DAVID
; APPLICANT: MANNING, GERARD
; APPLICANT: CAENEPEEL, SEAN
; TITLE OF INVENTION: NOVEL KINASES
; FILE REFERENCE: 034536-0321
; CURRENT APPLICATION NUMBER: US/10/618, 941
; PRIOR FILING DATE: 2003-07-15
; PRIOR APPLICATION NUMBER: 60/395, 632
; NUMBER OF SEQ ID NOS: 143
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 67
; LENGTH: 2055
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-618-941-67

Query Match 99.7%; Score 10464.5; DB 16; Length 2055;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2050; Conservative 2; Mismatches 2; Indels 1; Gaps 1;

Qy 1 MLKFKYGARNPLDAGAAEPRIASASRLNLFQGGKPPMTQQQMSPLSRBGLDALFVLF 60
Db 1 MLKFKYGARNPLDAGAAEPRIASASRLNLFQGGKPPMTQQQMSPLSRBGLDALFVLF 60
Qy 61 ECGOPALMKIKHNSNVPKRYSTPIAELOQASADFEVRSLVGGCHFEVQVVEKAT 119
Db 61 ECGOPALMKIKHNSNVPKRYSTPIAELOQASADFEVRSLVGGCHFEVQVVEKAT 120
Qy 120 GDYIANKVMKKALLAQEQVSPFEERBNILSRSTSPWIPOLQYAFODKNLVLVMEYOPG 179
Db 121 GDYIANKVMKKALLAQEQVSPFEERBNILSRSTSPWIPOLQYAFODKNLVLVMEYOPG 180
Qy 180 GDLLSLNRYEDQDENLQFYLAELILAVHSVHMGYVARDIKPENILVDRGHIKLV 239
Db 181 GDLLSLNRYEDQDENLQFYLAELILAVHSVHMGYVARDIKPENILVDRGHIKLV 240
Qy 240 FGSAAKKNNSKMNNAKLPTGTPYMAPEVLTWNNGGKGYGLDCMWSVGLAYEMTYG 299
Db 241 FGSAAKKNNSKMNNAKLPTGTPYMAPEVLTWNNGGKGYGLDCMWSVGLAYEMTYG 300
Qy 300 RSPFAEGTSARTFNINNFORFLKFPDDPKVSDFLDLIQSLCGQKERLKFEGLCCHPF 359
Db 301 RSPFAEGTSARTFNINNFORFLKFPDDPKVSDFLDLIQSLCGQKERLKFEGLCCHPF 360
Qy 360 FSKIDMNNINRNSPPFVPTLKSDDDTSNFDEPKNSWSSPCQLSPSGSGEELPFVGF 419
Db 361 FSKIDMNNINRNSPPFVPTLKSDDDTSNFDEPKNSWSSPCQLSPSGSGEELPFVGF 420
Qy 420 SYKALGILGRSSSVYGLDSPAKTSMKKLLIKSEKLOSDQCHKHKEQMTLHRVY 479
Db 421 SYKALGILGRSSSVYGLDSPAKTSMKKLLIKSEKLOSDQCHKHKEQMTLHRVY 480
Qy 480 SEVEAVLSQKEVELKASETORSLLEODLATYITECSSLKRSLEQARMEVSOEDKALQL 539
Db 481 SEVEAVLSQKEVELKASETORSLLEODLATYITECSSLKRSLEQARMEVSOEDKALQL 540
Qy 540 HDIREOSRKLQEIKEOEYQAOVEEMRLMNOULEEDLVASARRSDLYESELRESRLAAEF 599
Db 541 HDIREOSRKLQEIKEOEYQAOVEEMRLMNOULEEDLVASARRSDLYESELRESRLAAEF 600
Qy 600 KRYATBCQKLLAKAQOGKVEGEYAKLEKINAEQOLKIOELOKLEKAVKATETELL 659
Db 601 KRYATBCQKLLAKAQOGKVEGEYAKLEKINAEQOLKIOELOKLEKAVKATETELL 660
Qy 660 QNIRQAKERARELEKLONRDESSBGRKKLVABEERHSLKVKVRLTETMERERNLKD 719
Db 661 QNIRQAKERARELEKLONRDESSBGRKKLVABEERHSLKVKVRLTETMERERNLKD 720
Qy 720 DIQTKSQQLQOMADKILLEKEKREAOVSAQHLVHLKQEHYEKIKVLDNQIKKDLA 779
Db 721 DIQTKSQQLQOMADKILLEKEKREAOVSAQHLVHLKQEHYEKIKVLDNQIKKDLA 780
Qy 780 DKETLENMQRHEBEAHEKGIISBQKAMINANDSKIRSIHQIYIVELSEANKLAANSLSF 839
Db 781 DKETLENMQRHEBEAHEKGIISBQKAMINANDSKIRSIHQIYIVELSEANKLAANSLSF 840
Qy 840 TORNMKRAOEEMISELQOKFYLETQAGKLEAONRKLEBOLEKISHODSHDKNLTLETR 899
Db 841 TORNMKRAOEEMISELQOKFYLETQAGKLEAONRKLEBOLEKISHODSHDKNLTLETR 900
Qy 900 LREVSLEHEQKLEKRLQELTELQSLQERESQTLAQAAALAESOLQAKTELEETJAE 959
Db 901 LREVSLEHEQKLEKRLQELTELQSLQERESQTLAQAAALAESOLQAKTELEETJAE 960
Qy 960 ABEEIQALTAHDEIQKRPDALNSCTVITDLEQNLQUTEDNAEINNONFYLSKQIDEA 1019
Db 961 ABEEIQALTAHDEIQKRPDALNSCTVITDLEQNLQUTEDNAEINNONFYLSKQIDEA 1020
Qy 1020 SGANDEIVOLRSVDHLRREITEREMQUTSOKQTMALKTTCTMLEBOVNDLEALNDEL 1079
Db 1021 SGANDEIVOLRSVDHLRREITEREMQUTSOKQTMALKTTCTMLEBOVNDLEALNDEL 1080

1080 EKEROMEARSVLGDKEKQFECRVRELQRLDTEKQSRAPADORTTESROVVELAKREHK 1139
1081 EKEROMEARSVLGDKEKQFECRVRELQRLDTEKQSRAPADORTTESROVVELAKREHK 1140
1140 AEIILAQALKEOKLKASLSLSDKLNDEKKAJLEMMNRSLOQKLETEREKORLLEBOA 1199
1141 AEIILAQALKEOKLKASLSLSDKLNDEKKAJLEMMNRSLOQKLETEREKORLLEBOA 1200
1200 KLQOQMDLQKXHIIFRLTQGLQBALDRADLLKTERSDELEYOLENIQVLYSHEKVKMEGTIS 1259
1201 KLQOQMDLQKXHIIFRLTQGLQBALDRADLLKTERSDELEYOLENIQVLYSHEKVKMEGTIS 1260
1260 QOTKLIFLQAKMOPAKKKKVPLOQYNELKALEKAKCALEBEALOKTIELRSABEE 1319
1261 QOTKLIFLQAKMOPAKKKKVPLOQYNELKALEKAKCALEBEALOKTIELRSABEE 1320
1320 AAHRKATDHPSTPATRQOJAMSATVSRPEHOPSANSLAPSSRRKESSTPEEFRR 1379
1321 AAHRKATDHPSTPATRQOJAMSATVSRPEHOPSANSLAPSSRRKESSTPEEFRR 1380
1380 LKERMHNNIPHRFNVGLMRAATKCAVCLDTVHFGQASKLECYQMCHPKSTCLPATCG 1439
1381 LKERMHNNIPHRFNVGLMRAATKCAVCLDTVHFGQASKLECYQMCHPKSTCLPATCG 1440
1440 LPAEYATHTFAFCRDKNNSPGLQTEPSSSLHLEGMKVPRNNKRGQGDRIKXYILEG 1499
1441 LPAEYATHTFAFCRDKNNSPGLQTEPSSSLHLEGMKVPRNNKRGQGDRIKXYILEG 1500
1500 SKVLIIYDNEAREAGORPVEEFELCPDGDVSIHGAVGASELANTRAKDVPYILIMESHPH 1559
1501 SKVLIIYDNEAREAGORPVEEFELCPDGDVSIHGAVGASELANTRAKDVPYILIMESHPH 1560
1560 TTCWPGRTLYLLABSPDKORWVTALESVAVAGVSRKAKADAKLIGNSLKLLEGDRL 1619
1561 TTCWPGRTLYLLABSPDKORWVTALESVAVAGVSRKAKADAKLIGNSLKLLEGDRL 1620
1620 DMCNTLPPSDQVVLVGTREBGIYALNVLKNSLTHVPGIGAVQIYIIXDLEKLMIAEER 1679
1621 DMCNTLPPSDQVVLVGTREBGIYALNVLKNSLTHVPGIGAVQIYIIXDLEKLMIAEER 1680
1680 ALCLVDYKVKVQSLAQSHLPAPQDISPNIFFAVVGCFLFGAGKLENGLCICAAMPKRYI 1739
1681 ALCLVDYKVKVQSLAQSHLPAPQDISPNIFFAVVGCFLFGAGKLENGLCICAAMPKRYI 1740
1740 LRYNENLSTKCIKEIETSEPCSIHFTNYSILIGTNKFEYIDMKQYTLBEFLDKNDHSL 1799
1741 LRYNENLSTKCIKEIETSEPCSIHFTNYSILIGTNKFEYIDMKQYTLBEFLDKNDHSL 1800
1800 APAVFAASSNSFPVSIQVNSAGOREEYLLCFHEFGVFDVSYGRSRTDLMKGRPLAF 1859
1801 APAVFAASSNSFPVSIQVNSAGOREEYLLCFHEFGVFDVSYGRSRTDLMKGRPLAF 1860
1860 AYREPYLFVTHFNLELITELIOARSSAGTPARAAYLDINPRILGPAISSGAIYLLASSYQDK 1919
1861 AYREPYLFVTHFNLELITELIOARSSAGTPARAAYLDINPRILGPAISSGAIYLLASSYQDK 1920
1920 LRVYICCKGNLVKESGTEHHRPSTSRSSPNKRGPTNYNEHTTKVASSPAPPEGSHPRE 1979
1921 LRVYICCKGNLVKESGTEHHRPSTSRSSPNKRGPTNYNEHTTKVASSPAPPEGSHPRE 1980
1980 PSTPHRYREGRTELRDKSPGRPLEREKSPGRILSTRERSPARLFEDSSKGRLLPAGAVR 2039
1981 PSTPHRYREGRTELRDKSPGRPLEREKSPGRILSTRERSPARLFEDSSKGRLLPAGAVR 2040
2040 TPISQVNVKWDQSSV 2054
2041 TPISQVNVKWDQSSV 2055

RESULT 4
US-09-964-956-11
; Sequence 11, Application US/09964956
; Publication No. US20040043926A1

GENERAL INFORMATION:
APPLICANT: Gerlach, Valerie L
APPLICANT: MacDougall, John R
APPLICANT: Smithson, Glenda
APPLICANT: Millet, Isabelle
APPLICANT: Stone, David
APPLICANT: Gunther, Erik
APPLICANT: Ellerman, Karen
APPLICANT: Grose, William M
APPLICANT: Alsbrook II, John P
APPLICANT: Lepley, Denise M
APPLICANT: Burgess, Catherine E
APPLICANT: Padigaru, Muralidhara
APPLICANT: Kerkuda, Rameesh
APPLICANT: Spytek, Kimberly A
APPLICANT: Leach, Martin D
APPLICANT: Shinkets, Richard A
TITLE OF INVENTION: No. US20040043926A1 Proteins and Nucleic Acids Encoding Same
FILE REFERENCE: 21402-124
CURRENT APPLICATION NUMBER: US/09/964,956
CURRENT FILING DATE: 2001-09-26
PRIOR APPLICATION NUMBER: 60/235,631
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: 60/235,633
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: 60/235,808
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: 60/236,064
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: 60/236,065
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: 60/236,066
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: 60/236,135
PRIOR FILING DATE: 2000-09-28
PRIOR APPLICATION NUMBER: 60/237,434
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/238,321
PRIOR FILING DATE: 2000-10-05
PRIOR APPLICATION NUMBER: 60/238,399
PRIOR FILING DATE: 2000-10-06
PRIOR APPLICATION NUMBER: 60/238,396
PRIOR FILING DATE: 2000-10-06
PRIOR APPLICATION NUMBER: 60/276,667
PRIOR FILING DATE: 2001-03-16
PRIOR APPLICATION NUMBER: 60/294,823
PRIOR FILING DATE: 2001-05-31
PRIOR APPLICATION NUMBER: 60/304,868
PRIOR FILING DATE: 2001-07-12
NUMBER OF SEQ ID NOS: 127
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 11
LENGTH: 2053
TYPE: PRT
ORGANISM: Homo sapiens
US-09-964-956-11

Query Match 99.7%; Score 10458.5; DB 11; Length 2053;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2049; Conservative 2; Mismatches 2; Indels 1; Gaps 1;

1 MLKFKYGARPLDGAEPASIRASRLNLFQKGPPEMTQOQMSPLRREGILDLPLVLF 60
1 MLKFKYGARPLDGAEPASIRASRLNLFQKGPPEMTQOQMSPLRREGILDLPLVLF 60
61 ECGPALMKIKHVSNFVRKYSDDTIAELQPSAKDEEVSILVCGHFAEVQVVRKATG 120
61 ECGPALMKIKHVSNFVRKYSDDTIAELQPSAKDEEVSILVCGHFAEVQVVRKATG 120
121 DIYAMKMKKKALLAOQVSPFEERNTLSRSTSPMTPOIYAFODKMLYLVNRYPGG 180
121 DIYAMKMKKKALLAOQVSPFEERNTLSRSTSPMTPOIYAFODKMLYLVNRYPGG 180

QY 181 DLSTLNRYEDQJDENTLOFYLAELILAVSHVMGVHDIKRENTLNDRTGHIKLVDF 240
 DB 181 DLSTLNRYEDQJDENTLOFYLAELILAVSHVMGVHDIKRENTLNDRTGHIKLVDF 240
 QY 241 GSAAKNSNKNVNAKLPITGPDYMAPEVLTVMNGDGKTYGLJDCDMMVGVIAEMTYGR 300
 DB 241 GSAAKNSNKNVNAKLPITGPDYMAPEVLTVMNGDGKTYGLJDCDMMVGVIAEMTYGR 299
 QY 301 SPFAEGTSARTFNINMFORFLKPPDDPKVSSDFLDLIQSLICQOKERLKEGICCHPFF 360
 DB 301 SPFAEGTSARTFNINMFORFLKPPDDPKVSSDFLDLIQSLICQOKERLKEGICCHPFF 359
 QY 361 SKIDMNNIRSPPEFVTLKSDDDTSNPDEBEKNSWSSPCOLSPSGFSEELPFGFS 420
 DB 361 SKIDMNNIRSPPEFVTLKSDDDTSNPDEBEKNSWSSPCOLSPSGFSEELPFGFS 419
 QY 421 YSKALGILGRSESVVSGLDSPAKTSMCKLLIKSKELQDSQDKCHMEQEMTLHRVS 480
 DB 421 YSKALGILGRSESVVSGLDSPAKTSMCKLLIKSKELQDSQDKCHMEQEMTLHRVS 479
 QY 481 EBEAVLSQKEVELKASFTQSLBODLATYITTECSSLKRSLEQARMESQEDDQALQLH 540
 DB 481 EBEAVLSQKEVELKASFTQSLBODLATYITTECSSLKRSLEQARMESQEDDQALQLH 539
 QY 541 DIRQSKLOEIKQOEYQAOYEBEMLMNOLEBDLVASARRSDIYSEBELRESRLAAEFK 600
 DB 541 DIRQSKLOEIKQOEYQAOYEBEMLMNOLEBDLVASARRSDIYSEBELRESRLAAEFK 599
 QY 601 RKATECQHKLLKAKDQKPEVGEYAKLEKINAEOQLIOELQELBAVASTBATLLO 660
 DB 601 RKATECQHKLLKAKDQKPEVGEYAKLEKINAEOQLIOELQELBAVASTBATLLO 659
 QY 661 NIFQAKERARELEKLONREDSSGIRKQVVEABERHSHLENNVKLETFMERENRLKDD 720
 DB 661 NIFQAKERARELEKLONREDSSGIRKQVVEABERHSHLENNVKLETFMERENRLKDD 719
 QY 721 IQRKSOQIOQMADKILLEEKHREAOVSAOHLBTHLQKQOYBEKIKVLDNOIKKDLAD 780
 DB 721 IQRKSOQIOQMADKILLEEKHREAOVSAOHLBTHLQKQOYBEKIKVLDNOIKKDLAD 779
 QY 781 KETLENNQOHEEFAHKEGKILSEOKAMINAMDSKISLEORIVELSEANKLANSSFLT 840
 DB 781 KETLENNQOHEEFAHKEGKILSEOKAMINAMDSKISLEORIVELSEANKLANSSFLT 839
 QY 841 QRMNKAQEMISIELROQKFLYLETQAGKLEAONRKLBOLEKISHQDSDKNRLLEETRL 900
 DB 841 QRMNKAQEMISIELROQKFLYLETQAGKLEAONRKLBOLEKISHQDSDKNRLLEETRL 899
 QY 901 REVSLHEBQKLEKROLTELQSLQORRESQLTALQARAALBSQLROAKTELEETFAA 960
 DB 901 REVSLHEBQKLEKROLTELQSLQORRESQLTALQARAALBSQLROAKTELEETFAA 959
 QY 961 EBEIQTALTARDEIQRKFDALRNSCTVITDLEBQLOLQTEBNAELNNQNFYLSKQDDEAS 1020
 DB 961 EBEIQTALTARDEIQRKFDALRNSCTVITDLEBQLOLQTEBNAELNNQNFYLSKQDDEAS 1019
 QY 1021 GANDEIVOLSEVDHLRREITEREMOULTSQOTWEALKTCTMLEBQVMDLEALNDELLE 1080
 DB 1021 GANDEIVOLSEVDHLRREITEREMOULTSQOTWEALKTCTMLEBQVMDLEALNDELLE 1079
 QY 1081 KERQWEMRVSVLGDEKQOFERVALEORMLDTEKQSPARADORTTEROYVELAVKHNKA 1140
 DB 1081 KERQWEMRVSVLGDEKQOFERVALEORMLDTEKQSPARADORTTEROYVELAVKHNKA 1139
 QY 1141 EILALQALKEOKUKASLSQDKLNDLEKQAMLENNARSLQOKLETREBELKORLLEBQAK 1200
 DB 1141 EILALQALKEOKUKASLSQDKLNDLEKQAMLENNARSLQOKLETREBELKORLLEBQAK 1199
 QY 1201 LQOQMDLOKNHIFRLTGLOALDRADULKTERSDLEYOLENTQVLYSHEKVNKEGTISQ 1260
 DB 1201 LQOQMDLOKNHIFRLTGLOALDRADULKTERSDLEYOLENTQVLYSHEKVNKEGTISQ 1259
 QY 1261 QTKLIDFLQAKMDQPAKKKVPLQYNELKALEKKAACALEBALQKTIIELSAREEA 1320

DB 1260 QTKLIDFLQAKMDQPAKKKVPLQYNELKALEKKAACALEBALQKTIIELSAREEA 1319
 QY 1321 ABRKATDHPSPSTPATARQOIAMSAIVRSEBOPSAMSLIAPSSSRKESSTPEEFSRRL 1380
 DB 1320 ABRKATDHPSPSTPATARQOIAMSAIVRSEBOPSAMSLIAPSSSRKESSTPEEFSRRL 1379
 QY 1381 KERMHNNIPHRFVNGIMNARTKCAVCLDPTVHFORQASKLECOVMCHPKSTCLPNTCG 1440
 DB 1380 KERMHNNIPHRFVNGIMNARTKCAVCLDPTVHFORQASKLECOVMCHPKSTCLPNTCG 1439
 QY 1441 PAEYATHTEAPCRDKMNSPGLQTKPESSILHEGMMKVPNNKRGQOQWDRKYIVLEGS 1500
 DB 1440 PAEYATHTEAPCRDKMNSPGLQTKPESSILHEGMMKVPNNKRGQOQWDRKYIVLEGS 1499
 QY 1501 KVLIDYNEAREAQORPVEEFELCLPGDVSIHGAVQASELANATAKADVPIYILMESHPHT 1560
 DB 1500 KVLIDYNEAREAQORPVEEFELCLPGDVSIHGAVQASELANATAKADVPIYILMESHPHT 1559
 QY 1561 TCMFGRTLYILASPFDKQWMTALSVVAGGVSRKKAADAKLIGNSILKLEGGDRLD 1620
 DB 1560 TCMFGRTLYILASPFDKQWMTALSVVAGGVSRKKAADAKLIGNSILKLEGGDRLD 1619
 QY 1621 MNCPLPSPDOVNVGTEBGLYALNVLKNSLTHVPGIGAVFOIYIIXDLKELMAGEERA 1680
 DB 1620 MNCPLPSPDOVNVGTEBGLYALNVLKNSLTHVPGIGAVFOIYIIXDLKELMAGEERA 1679
 QY 1681 LCLVDYKVKQSLAQSHLPAOPDISPNIFEAVKGCILFGAKIENGICICAMPKRVIL 1740
 DB 1680 LCLVDYKVKQSLAQSHLPAOPDISPNIFEAVKGCILFGAKIENGICICAMPKRVIL 1739
 QY 1741 RYNEBNSKTCIRKEITSEPCSCIHFTNYSILIGTKKFEYIMDKOYTLBEFDKXNHSIA 1800
 DB 1740 RYNEBNSKTCIRKEITSEPCSCIHFTNYSILIGTKKFEYIMDKOYTLBEFDKXNHSIA 1799
 QY 1801 PAVFAASNSFPVSIYQVNSAGOREEYLCFHEFGVFDVSYGRSRTDLMKRLPLAFA 1860
 DB 1800 PAVFAASNSFPVSIYQVNSAGOREEYLCFHEFGVFDVSYGRSRTDLMKRLPLAFA 1859
 QY 1861 YREPYLVTHFNSLEVEIEIQARSSAGTPARAVLDINPRYLGPAISSGAIYLAASYODKL 1920
 DB 1860 YREPYLVTHFNSLEVEIEIQARSSAGTPARAVLDINPRYLGPAISSGAIYLAASYODKL 1919
 QY 1921 RVICCKGNLVKESGTEHHRGPTSRSSPNKRGPTYNENHTTKVASSPAPBEPSPHREP 1979
 DB 1920 RVICCKGNLVKESGTEHHRGPTSRSSPNKRGPTYNENHTTKVASSPAPBEPSPHREP 1979
 QY 1981 STPHRYREGTELRKDKSPGRPLEREKSPGRILSTRERSPARLPEDSSRGRLPAGAVRT 2040
 DB 1980 STPHRYREGTELRKDKSPGRPLEREKSPGRILSTRERSPARLPEDSSRGRLPAGAVRT 2039
 QY 2041 PLSQVNVKMDQSSV 2054
 DB 2040 PLSQVNVKMDQSSV 2053

RESULT 5
 US-10-262-511-2
 ; Sequence 2, Application US/10262511
 ; Publication No. US20040038223A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Smithsonian, Glenda
 ; APPLICANT: Millet, Isabelle
 ; APPLICANT: Peyman, John A.
 ; APPLICANT: Kerkuda, Ramesh
 ; APPLICANT: Ju, Jinfang
 ; APPLICANT: Li, Li
 ; APPLICANT: Guo, Xiaojia (Sasha)
 ; APPLICANT: Patuturajan, Meera
 ; APPLICANT: Spytek, Kimberly A.
 ; APPLICANT: Edinger, Shlomit R.
 ; APPLICANT: Ellerman, Karen
 ; APPLICANT: Malysankar, Uriel M.

```

? APPLICANT: Ort, Tatiana
? APPLICANT: Gorman, Linda
? APPLICANT: Zernusen, Bryan D.
? APPLICANT: Anderson, David W.
? APPLICANT: Zhong, Mei
? APPLICANT: Carterton, Elina
? APPLICANT: J. Weizhen
? APPLICANT: Miller, Charles E.
? APPLICANT: Raetelli, Luca
? APPLICANT: Stone, David J.
? APPLICANT: Pena, Carol E. A.
? APPLICANT: Shenoy, Suresh G.
? APPLICANT: Shimkets, Richard A.
? APPLICANT: Rothenberg, Mark E.
? APPLICANT: Leach, Martin D.
? APPLICANT: Agee, Michele L.
? APPLICANT: Berghs, Constance
? TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
? FILE REFERENCE: 21402-462C
? CURRENT APPLICATION NUMBER: US/10/262,511
? PRIOR APPLICATION NUMBER: 2003-05-28
? PRIOR FILING DATE: 2001-10-02
? PRIOR APPLICATION NUMBER: 60/326,483
? PRIOR FILING DATE: 2002-04-19
? PRIOR APPLICATION NUMBER: 60/373,815
? PRIOR FILING DATE: 2001-10-09
? PRIOR APPLICATION NUMBER: 60/327,917
? PRIOR FILING DATE: 2001-10-09
? PRIOR APPLICATION NUMBER: 60/381,642
? PRIOR FILING DATE: 2002-05-17
? PRIOR APPLICATION NUMBER: 60/328,029
? PRIOR FILING DATE: 2002-10-09
? PRIOR APPLICATION NUMBER: 60/381,038
? PRIOR FILING DATE: 2002-05-16
? PRIOR APPLICATION NUMBER: 60/328,056
? PRIOR FILING DATE: 2001-10-09
? PRIOR APPLICATION NUMBER: 60/373,260
? PRIOR FILING DATE: 2002-04-17
? PRIOR APPLICATION NUMBER: 60/373,826
? PRIOR FILING DATE: 2002-04-19
? PRIOR APPLICATION NUMBER: 60/327,435
? PRIOR FILING DATE: 2001-10-05
? Remaining Prior Application data removed - See File Wrapper or PALM.
? NUMBER OF SEQ ID NOS: 439
? SOFTWARE: Curoseq1ist version 0.1
? SEQ ID NO 2
? LENGTH: 2053
? TYPE: PRF
? ORGANISM: Homo sapiens
US-10-262-511-2

Query Match      99.7%; Score 10458.5; DB 15; Length 2053;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2049; Conservative 2; Mismatches 2; Indels 1; Gaps 1;

```

```

QY 301 SPFAEGTSARTFNNINPQRFLEKPPDDPKVSSDFLDLIQSLLCGQKERLEKGLCCHPFF 360
DB 300 SPFAEGTSARTFNNINPQRFLEKPPDDPKVSSDFLDLIQSLLCGQKERLEKGLCCHPFF 359
QY 361 SKIDMNNIRNSPPFPVFTLSDDDTSNFDPEBKNSWSSSPQCQSPSGFSSEELPVPVGS 420
DB 360 SKIDMNNIRNAPFPFVFTLSDDDTSNFDPEBKNSWSSSPQCQSPSGFSSEELPVPVGS 419
QY 421 YSKALGILGRSESVSGSDSPAKTSMEKLLISKELODSQDCHMKMEQMTLHRVVS 480
DB 420 YSKALGILGRSESVSGSDSPAKTSMEKLLISKELODSQDCHMKMEQMTLHRVVS 479
QY 481 EVEAVLSQKEVELKASTORSLLBODLATYITTECSSIKRSLSQARMVEVSQBDKALQLLH 540
DB 480 EVEAVLSQKEVELKASTORSLLBODLATYITTECSSIKRSLSQARMVEVSQBDKALQLLH 539
QY 541 DIRQSRKLOEIKQOEYQAOVEEMRLMNNQLEBVLVARRRSDLYESBLRSLAAEFK 600
DB 540 DIRQSRKLOEIKQOEYQAOVEEMRLMNNQLEBVLVARRRSDLYESBLRSLAAEFK 599
QY 601 RKATECOHKLKAKDOGKPEVGEYAKLEKINAEOQLKIOELOEKLKAVKASTATEATLQ 660
DB 600 RKATECOHKLKAKDOGKPEVGEYAKLEKINAEOQLKIOELOEKLKAVKASTATEATLQ 659
QY 661 NIRQAKERARELEKLNREDSEGIKKULVEAEERHSHLENKYKRLTERRERENRLKDD 720
DB 660 NIRQAKERARELEKLNREDSEGIKKULVEAEERHSHLENKYKRLTERRERENRLKDD 719
QY 721 IQTSQOIQOMADKILBELKREHQAQVSAQHLVHLKQKQHYBEKIKVLDNOIKKDLAD 780
DB 720 IQTSQOIQOMADKILBELKREHQAQVSAQHLVHLKQKQHYBEKIKVLDNOIKKDLAD 779
QY 781 KETLENNMOHBEERAEHKGKILSEOKAMINAMSKIRSEORIVELSEANKLANSSLFT 840
DB 780 KETLENNMOHBEERAEHKGKILSEOKAMINAMSKIRSEORIVELSEANKLANSSLFT 839
QY 841 QRNMKAQEMISIELRQKFLYLETQAGLEAONRKLBEOLKISHQDSDKXRLLELETRL 900
DB 840 QRNMKAQEMISIELRQKFLYLETQAGLEAONRKLBEOLKISHQDSDKXRLLELETRL 899
QY 901 REVSLEHEBOQLBIKROLTELQUSLORESQTLNQAARALAESQLRQATELETTAAEA 960
DB 900 REVSLEHEBOQLBIKROLTELQUSLORESQTLNQAARALAESQLRQATELETTAAEA 959
QY 961 EEEIQALTARDEIQRFDAIRNSCTVITDLEBQNLTEDMELNNQNFYLSQLEBAS 1020
DB 960 EEEIQALTARDEIQRFDAIRNSCTVITDLEBQNLTEDMELNNQNFYLSQLEBAS 1019
QY 1021 GANDEIVQLRSEVDHLREITTEREMOLTSQKQTEALKTTCTMLEBQVMDLEALNDEILE 1080
DB 1020 GANDEIVQLRSEVDHLREITTEREMOLTSQKQTEALKTTCTMLEBQVMDLEALNDEILE 1079
QY 1081 KERQWEAMRSVLGDEKSQFECRAVREIQRMUTDKOSPARADORTESROVVEALVKEHKA 1140
DB 1080 KERQWEAMRSVLGDEKSQFECRAVREIQRMUTDKOSPARADORTESROVVEALVKEHKA 1139
QY 1141 EIIALQALKEOKLKAKESLSDKLNDEKKGAMLEMANRSIQKLETERELKORLLEBOAK 1200
DB 1140 EIIALQALKEOKLKAKESLSDKLNDEKKGAMLEMANRSIQKLETERELKORLLEBOAK 1199
QY 1201 LQQQMDLQKNHIFRLTQGLQBALDRADLLKTERSSDLEYQLENIQVLSHSEKVKMEGTSIQ 1260
DB 1200 LQQQMDLQKNHIFRLTQGLQBALDRADLLKTERSSDLEYQLENIQVLSHSEKVKMEGTSIQ 1259
QY 1261 QTKLIDFLQAKMDQPAKKKVPLOYNELKLALKEKARCALEBALOKTRIIELSAREEA 1320
DB 1260 QTKLIDFLQAKMDQPAKKKVPLOYNELKLALKEKARCALEBALOKTRIIELSAREEA 1319
QY 1321 AHRKATDHPHSPPTATPQOIIASATVSRSPHQPSAMSLAPSSSRKESSTPEEFRRLL 1380
DB 1320 AHRKATDHPHSPPTATPQOIIASATVSRSPHQPSAMSLAPSSSRKESSTPEEFRRLL 1379

```

QY 1381 KERMHNIPRRVNGVIMRATKCAVCLDTVHFGROASKLEQVMCHPKGSTCLPATCGL 1440
DB 1380 KERMHNIPRRVNGVIMRATKCAVCLDTVHFGROASKLEQVMCHPKGSTCLPATCGL 1439
QY 1441 PAVYATHTFAFGCDKMSNPGLOTKEPSSILHLEGMMKVRNNKRGQGDRIKTYIVLEGS 1500
DB 1440 PAVYATHTFAFGCDKMSNPGLOTKEPSSILHLEGMMKVRNNKRGQGDRIKTYIVLEGS 1499
QY 1501 KVLIDNEAREAGQRPVEEFELCLPDGDVSIHGAVGASELANAKADVPYILKMEHPHT 1560
DB 1500 KVLIDNEAREAGQRPVEEFELCLPDGDVSIHGAVGASELANAKADVPYILKMEHPHT 1559
QY 1561 TCMRGRTLYLLASFPKQWMTALBSVAVAGRRSPRKADAKLGNLSILXLEGDRD 1620
DB 1560 TCMRGRTLYLLASFPKQWMTALBSVAVAGRRSPRKADAKLGNLSILXLEGDRD 1619
QY 1621 MNCITLPPSDQVVLVGTREEGLYALNVLNKSILTHVIGIAVEQIYIILKDEKLMIAGERA 1680
DB 1620 MNCITLPPSDQVVLVGTREEGLYALNVLNKSILTHVIGIAVEQIYIILKDEKLMIAGERA 1679
QY 1681 LCLVDVKVKQSLAQSHLPAQPDISPNIFFAVKGCILFGAKILENGLCICAMPKSVYIL 1740
DB 1680 LCLVDVKVKQSLAQSHLPAQPDISPNIFFAVKGCILFGAKILENGLCICAMPKSVYIL 1739
QY 1741 RYNNENSKYCIKREIETSEPCSCIHFTNYSILIGTKKFEYIIMKQYTLSEFIDKNDHSLA 1800
DB 1740 RYNNENSKYCIKREIETSEPCSCIHFTNYSILIGTKKFEYIIMKQYTLSEFIDKNDHSLA 1799
QY 1801 PAVFAASNSFPYSIVOVNSAGOREEYLLCFHEGTVDSYGRSRRTDDKMSRLPLAFA 1860
DB 1800 PAVFAASNSFPYSIVOVNSAGOREEYLLCFHEGTVDSYGRSRRTDDKMSRLPLAFA 1859
QY 1861 YREPYLVFTHFNSLEVLIEIQARSSAGTPARAYLDIPNRYLGAIISSGAIYLAASYDCL 1920
DB 1860 YREPYLVFTHFNSLEVLIEIQARSSAGTPARAYLDIPNRYLGAIISSGAIYLAASYDCL 1919
QY 1921 RYVCCGNLVKESGTEHHRPSTSRSSPNKRGPTTYNEHTTKVASSPAPEGSHPREP 1980
DB 1920 RYVCCGNLVKESGTEHHRPSTSRSSPNKRGPTTYNEHTTKVASSPAPEGSHPREP 1979
QY 1981 STEHRYEGRTELRLRDSPPRLEREKSPGRIILSTRERSPARLLFEDSSRGRLPAGAVRT 2040
DB 1980 STEHRYEGRTELRLRDSPPRLEREKSPGRIILSTRERSPARLLFEDSSRGRLPAGAVRT 2039
QY 2041 PLSQVNVKWDQSSV 2054
DB 2040 PLSQVNVKWDQSSV 2053

RESULT 6
US-09-964-956-9
; Sequence 9, Application US/09964956
; Publication No. US20040043926A1
; GENERAL INFORMATION:
; APPLICANT: Gerlach, Valerie L
; APPLICANT: Macdougall, John R
; APPLICANT: Smitheon, Glenda
; APPLICANT: Miller, Isabelle
; APPLICANT: Stone, David
; APPLICANT: Gunther, Erik
; APPLICANT: Ellerman, Karen
; APPLICANT: Grosche, William M
; APPLICANT: Alsobrook II, John P
; APPLICANT: Lepley, Denise M
; APPLICANT: Burgess, Catherine B
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Kekuda, Rameen
; APPLICANT: Spytek, Kimberly A
; APPLICANT: Leach, Martin D
; APPLICANT: Shimkets, Richard A
; TITLE OF INVENTION: No. US20040043926A1
; FILE REFERENCE: 21402-124
; CURRENT APPLICATION NUMBER: US/09/964, 956

QY 1381 KERMHNIPRRVNGVIMRATKCAVCLDTVHFGROASKLEQVMCHPKGSTCLPATCGL 1440
DB 1380 KERMHNIPRRVNGVIMRATKCAVCLDTVHFGROASKLEQVMCHPKGSTCLPATCGL 1439
QY 1441 PAVYATHTFAFGCDKMSNPGLOTKEPSSILHLEGMMKVRNNKRGQGDRIKTYIVLEGS 1500
DB 1440 PAVYATHTFAFGCDKMSNPGLOTKEPSSILHLEGMMKVRNNKRGQGDRIKTYIVLEGS 1499
QY 1501 KVLIDNEAREAGQRPVEEFELCLPDGDVSIHGAVGASELANAKADVPYILKMEHPHT 1560
DB 1500 KVLIDNEAREAGQRPVEEFELCLPDGDVSIHGAVGASELANAKADVPYILKMEHPHT 1559
QY 1561 TCMRGRTLYLLASFPKQWMTALBSVAVAGRRSPRKADAKLGNLSILXLEGDRD 1620
DB 1560 TCMRGRTLYLLASFPKQWMTALBSVAVAGRRSPRKADAKLGNLSILXLEGDRD 1619
QY 1621 MNCITLPPSDQVVLVGTREEGLYALNVLNKSILTHVIGIAVEQIYIILKDEKLMIAGERA 1680
DB 1620 MNCITLPPSDQVVLVGTREEGLYALNVLNKSILTHVIGIAVEQIYIILKDEKLMIAGERA 1679
QY 1681 LCLVDVKVKQSLAQSHLPAQPDISPNIFFAVKGCILFGAKILENGLCICAMPKSVYIL 1740
DB 1680 LCLVDVKVKQSLAQSHLPAQPDISPNIFFAVKGCILFGAKILENGLCICAMPKSVYIL 1739
QY 1741 RYNNENSKYCIKREIETSEPCSCIHFTNYSILIGTKKFEYIIMKQYTLSEFIDKNDHSLA 1800
DB 1740 RYNNENSKYCIKREIETSEPCSCIHFTNYSILIGTKKFEYIIMKQYTLSEFIDKNDHSLA 1799
QY 1801 PAVFAASNSFPYSIVOVNSAGOREEYLLCFHEGTVDSYGRSRRTDDKMSRLPLAFA 1860
DB 1800 PAVFAASNSFPYSIVOVNSAGOREEYLLCFHEGTVDSYGRSRRTDDKMSRLPLAFA 1859
QY 1861 YREPYLVFTHFNSLEVLIEIQARSSAGTPARAYLDIPNRYLGAIISSGAIYLAASYDCL 1920
DB 1860 YREPYLVFTHFNSLEVLIEIQARSSAGTPARAYLDIPNRYLGAIISSGAIYLAASYDCL 1919
QY 1921 RYVCCGNLVKESGTEHHRPSTSRSSPNKRGPTTYNEHTTKVASSPAPEGSHPREP 1980
DB 1920 RYVCCGNLVKESGTEHHRPSTSRSSPNKRGPTTYNEHTTKVASSPAPEGSHPREP 1979
QY 1981 STEHRYEGRTELRLRDSPPRLEREKSPGRIILSTRERSPARLLFEDSSRGRLPAGAVRT 2040
DB 1980 STEHRYEGRTELRLRDSPPRLEREKSPGRIILSTRERSPARLLFEDSSRGRLPAGAVRT 2039
QY 2041 PLSQVNVKWDQSSV 2054
DB 2040 PLSQVNVKWDQSSV 2053

Query Match 99.4%; Score 10425.5; DB 11; Length 2066;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 2044; Conservative 2; Mismatches 5; Indels 1; Gaps 1;
US-09-964-956-9

QY 1 MLEKFGANPDLGAGAAPIASRASRLNLFQGGPPEMTQQQMSPLSREGILDALFVLF 60
DB 1 MLEKFGANPDLGAGAAPIASRASRLNLFQGGPPEMTQQQMSPLSREGILDALFVLF 60
QY 61 ECGOPALMKIKHVSNFVRKXSDTIALQELQPSAKDFEYRSLVGCGHFAEVQVYREKATG 120
DB 61 ECGOPALMKIKHVSNFVRKXSDTIALQELQPSAKDFEYRSLVGCGHFAEVQVYREKATG 120
QY 121 DIYAMKVMKKKLLAEOVSFFEEERNILSRSTSPMIPOLQVAFODKHLTYVMEXOPG 180
DB 121 DIYAMKVMKKKLLAEOVSFFEEERNILSRSTSPMIPOLQVAFODKHLTYVMEXOPG 180
QY 181 DLSTLNRYEDQDENLQFYLAELILAVSVHLMGVYHRDIPENILVDRTHIKLVDF 240
DB 181 DLSTLNRYEDQDENLQFYLAELILAVSVHLMGVYHRDIPENILVDRTHIKLVDF 240
QY 241 GSAAKNSNMVNAKLPITGPDYMADEVLTVNMGDKGTGYLDCDWMASGVIAIYEMIYGR 300
DB 241 GSAAKNSNMVNAKLPITGPDYMADEVLTVNMGDKGTGYLDCDWMASGVIAIYEMIYGR 299
QY 301 SPFAEGTSARTFNINNFQRLKFPDDPKVSSPFLDILQSLCGQERLKFEGLCCHPFF 360
DB 301 SPFAEGTSARTFNINNFQRLKFPDDPKVSSPFLDILQSLCGQERLKFEGLCCHPFF 359
QY 360 SKIDMNNIRNPPFPYTLKSDDTNSFDEPKNSVWSSPCQLSPSGFGEELPFVGF 420
DB 360 SKIDMNNIRNPPFPYTLKSDDTNSFDEPKNSVWSSPCQLSPSGFGEELPFVGF 419
QY 421 YSRALGILGRSESVGLDSPATYSMEKKLLIKSRELQDSQDKHMEQEMTRLHRRVS 480
DB 420 YSRALGILGRSESVGLDSPATYSMEKKLLIKSRELQDSQDKHMEQEMTRLHRRVS 479

QY 481 EVEAVLSQKEVELKASFTQBSLLBODLATYITTECSLSKRSLEQARMEVSYQSDRXLQJLH 540
 DB 480 EVEAVLSQKEVELKASFTQBSLLBODLATYITTECSLSKRSLEQARMEVSYQSDRXLQJLH 539
 QY 541 DIREQSRLQEIKEOEYQAOVEEMLMNNOLEEDVARSRRSDYSESLEBSRLAAEFK 600
 DB 540 DIREQSRLQEIKEOEYQAOVEEMLMNNOLEEDVARSRRSDYSESLEBSRLAAEFK 599
 QY 601 RKATECOHKLKAKDOQKPEVGEYAKLEKINAEQOLKIQELQEKLEKAVKASTETELLQ 660
 DB 600 RKATECOHKLKAKDOQKPEVGEYAKLEKINAEQOLKIQELQEKLEKAVKASTETELLQ 659
 QY 661 NIRAQKAEARELEKLONRBDSSEIGIRKLVABERRRSLLENKRYLETMERREKRLKOD 720
 DB 660 NIRAQKAEARELEKLONRBDSSEIGIRKLVABERRRSLLENKRYLETMERREKRLKOD 719
 QY 721 IOTSQOIQOMADKILBLEEKHREAOVSAOHLVYLKQEQHYEKKIVLNOIKODLAD 780
 DB 720 IOTSQOIQOMADKILBLEEKHREAOVSAOHLVYLKQEQHYEKKIVLNOIKODLAD 779
 QY 781 KETLENNMQRHEEAEHEKGIILSEOKAMINAMDSKIRSLBORIVELSEANKLAANSILFT 840
 DB 780 KETLENNMQRHEEAEHEKGIILSEOKAMINAMDSKIRSLBORIVELSEANKLAANSILFT 839
 QY 841 QRNKKAQEBMTSELRQCKFYLETQAGKLEAONRKLEBOLKISHODSDKRRLLBLETRL 900
 DB 840 QRNKKAQEBMTSELRQCKFYLETQAGKLEAONRKLEBOLKISHODSDKRRLLBLETRL 899
 QY 901 REVSLEHEEOKLEIKROLTTELQSLQEHESQTLQAARALEBOLRAKTELEETTAEA 960
 DB 900 REVSLEHEEOKLEIKROLTTELQSLQEHESQTLQAARALEBOLRAKTELEETTAEA 959
 QY 961 EEEIOALTARHDEIQRFKDALRNSCTVITDLEEOINQUTEDNABLANQNFYLSKOLDEAS 1020
 DB 960 EEEIOALTARHDEIQRFKDALRNSCTVITDLEEOINQUTEDNABLANQNFYLSKOLDEAS 1019
 QY 1021 GANDEIVOLRSEVDHLRREITEREMOLTQOKOTBEALKTCTMLEBOYMDLEALNDELLE 1080
 DB 1020 GANDEIVOLRSEVDHLRREITEREMOLTQOKOTBEALKTCTMLEBOYMDLEALNDELLE 1079
 QY 1081 KERQWEMARSVLGDKEGQFECRVRLEQOMLDEKOSRRARORITSESQVVELVKEKA 1140
 DB 1080 KERQWEMARSVLGDKEGQFECRVRLEQOMLDEKOSRRARORITSESQVVELVKEKA 1139
 QY 1141 EILALQOALKEQKKAESLSKNDLEKGMLENNASLQOKLETERELKORLLEBOAK 1200
 DB 1140 EILALQOALKEQKKAESLSKNDLEKGMLENNASLQOKLETERELKORLLEBOAK 1199
 QY 1201 LQQQMDLQKNHIFRLTQGLQBALDRADLLKTERSDBLEYOLENIQVLSHEKVKMEGTISQ 1260
 DB 1200 LQQQMDLQKNHIFRLTQGLQBALDRADLLKTERSDBLEYOLENIQVLSHEKVKMEGTISQ 1259
 QY 1261 QTKLIDFLQAMQDPACKKKVPLQYNELKALEKEKARCABLEBALQKTRTELKSAREEA 1320
 DB 1260 QTKLIDFLQAMQDPACKKKVPLQYNELKALEKEKARCABLEBALQKTRTELKSAREEA 1319
 QY 1321 AHRKATDHPHSTPATARQOJAMSALVRSPEHOPSAMSILAPSSRRKESSTPEEFSRL 1380
 DB 1320 AHRKATDHPHSTPATARQOJAMSALVRSPEHOPSAMSILAPSSRRKESSTPEEFSRL 1379
 QY 1381 KERMHNI PHRFNVGLNMRATKCAVCLDTVHFGROASKCLECOVMCHCKSTCLPATCGL 1440
 DB 1380 KERMHNI PHRFNVGLNMRATKCAVCLDTVHFGROASKCLECOVMCHCKSTCLPATCGL 1439
 QY 1441 PAVATHTPEAFCDKKNNSPGLQTKPESSSLHLSGMKVPNNKRGQOGQMDRKITYLEGS 1500
 DB 1440 PAVATHTPEAFCDKKNNSPGLQTKPESSSLHLSGMKVPNNKRGQOGQMDRKITYLEGS 1499
 QY 1501 KVLVYDNEARAGORPYEEFELCPDGDVSIHGAVGASELNTAKADVPIYLKIMESHPT 1560
 DB 1500 KVLVYDNEARAGORPYEEFELCPDGDVSIHGAVGASELNTAKADVPIYLKIMESHPT 1559
 QY 1561 TCWPGRTLTYLLAPSPDKQRWVTALBSVAVAGRVSRKAEADAKLLGNSLLKLEGGDRLD 1620

DB 1660 TCWPGRTLTYLLAPSPDKQRWVTALBSVAVAGRVSRKAEADAKLLGNSLLKLEGGDRLD 1619
 QY 1621 MNCITLPSDDQVVLVGTBEGLYALNVLKNSLTHVNGIGAVFOIYIIXKDEKLMIMAGEERA 1680
 DB 1620 MNCITLPSDDQVVLVGTBEGLYALNVLKNSLTHVNGIGAVFOIYIIXKDEKLMIMAGEERA 1679
 QY 1681 LCLVDVKKVKOSLAQSHLPAQPDISPNIPEAVKGCHEFGAKIENGCLCTCAMPKSVYL 1740
 DB 1680 LCLVDVKKVKOSLAQSHLPAQPDISPNIPEAVKGCHEFGAKIENGCLCTCAMPKSVYL 1739
 QY 1741 RYNNELSKYCTIKRKEITFSEPCSHFTMYSLILIGNKRYEIDMQYTLIEPLDKNDHSLA 1800
 DB 1740 RYNNELSKYCTIKRKEITFSEPCSHFTMYSLILIGNKRYEIDMQYTLIEPLDKNDHSLA 1799
 QY 1801 PAVFAASNSFPVSVIVQVNSAGOREBYLLCFHEFGVFPVDSIGRRSRDTDLKMSRLPLAFA 1860
 DB 1800 PAVFAASNSFPVSVIVQVNSAGOREBYLLCFHEFGVFPVDSIGRRSRDTDLKMSRLPLAFA 1859
 QY 1861 YREPYLFTVTHNSLEVIEIQARSSAGTPARAVIDIPNRYLGPALISSGAIYLAASYODKL 1920
 DB 1860 YREPYLFTVTHNSLEVIEIQARSSAGTPARAVIDIPNRYLGPALISSGAIYLAASYODKL 1919
 QY 1921 RVICCKGNLVKESGTEHHRRGSTSRSSPNKRGPTTYNEHTKRVASSAPPEGSHPREP 1980
 DB 1920 RVICCKGNLVKESGTEHHRRGSTSRSSPNKRGPTTYNEHTKRVASSAPPEGSHPREP 1979
 QY 1981 STPHRYREGRTELRRDSSPGRLERKSPGRILSTRRERSPARLPEDSSRRLPLGAART 2040
 DB 1980 STPHRYREGRTELRRDSSPGRLERKSPGRILSTRRERSPARLPEDSSRRLPLGAART 2039
 QY 2041 PLSQVNNKRVQDS 2052
 DB 2040 PLSQVNNKRVQDS 2051

RESULT 7
 US-10-262-511-14
 ; Sequence 14, Application US/10262511
 ; Publication No. US20040038223A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Smithson, Glenda
 ; APPLICANT: Miller, Isabelle
 ; APPLICANT: Peyman, John A.
 ; APPLICANT: Kekuda, Rameesh
 ; APPLICANT: Ju, Jingfang
 ; APPLICANT: Li, Li
 ; APPLICANT: Guo, Xiaojia (Sasha)
 ; APPLICANT: Paturajan, Meera
 ; APPLICANT: Spytek, Kimberly A.
 ; APPLICANT: Edinger, Shlomit R.
 ; APPLICANT: Elleman, Karen
 ; APPLICANT: Malyankar, Uriel M.
 ; APPLICANT: Ort, Tatiana
 ; APPLICANT: Zernusen, Bryan D.
 ; APPLICANT: Anderson, David W.
 ; APPLICANT: Zhong, Mei
 ; APPLICANT: Catterton, Elna
 ; APPLICANT: Ji, Weizhen
 ; APPLICANT: Miller, Charles E.
 ; APPLICANT: Raetelli, Luca
 ; APPLICANT: Stone, David J.
 ; APPLICANT: Pena, Carol E. A.
 ; APPLICANT: Shenoy, Suresh G.
 ; APPLICANT: Shimkets, Richard A.
 ; APPLICANT: Rothenberg, Mark E.
 ; APPLICANT: Leach, Martin D.
 ; APPLICANT: Agee, Michele L.
 ; APPLICANT: Berghs, Constance
 ; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
 ; FILE REFERENCE: 21402-462C
 ; CURRENT APPLICATION NUMBER: US/10/262,511

CURRENT FILING DATE: 2003-05-28
PRIOR APPLICATION NUMBER: 60/326,483
PRIOR FILING DATE: 2001-10-02
PRIOR APPLICATION NUMBER: 60/373,815
PRIOR FILING DATE: 2002-04-19
PRIOR APPLICATION NUMBER: 60/327,917
PRIOR FILING DATE: 2001-10-09
PRIOR APPLICATION NUMBER: 60/381,642
PRIOR FILING DATE: 2002-05-17
PRIOR APPLICATION NUMBER: 60/328,029
PRIOR FILING DATE: 2002-10-09
PRIOR APPLICATION NUMBER: 60/381,038
PRIOR FILING DATE: 2002-05-16
PRIOR APPLICATION NUMBER: 60/328,056
PRIOR FILING DATE: 2001-10-09
PRIOR APPLICATION NUMBER: 60/373,260
PRIOR FILING DATE: 2002-04-17
PRIOR APPLICATION NUMBER: 60/373,826
PRIOR FILING DATE: 2002-04-19
PRIOR APPLICATION NUMBER: 60/327,435
PRIOR FILING DATE: 2001-10-05
Remaining Prior Application data removed - See file wrapper or PALM.
NUMBER OF SEQ ID NOS: 439
SOFTWARE: Cuiaseqdist version 0.1
SEQ ID NO 14
LENGTH: 2066
TYPE: PRT
ORGANISM: Homo sapiens
US-10-262-511-14

Query Match 99.4%; Score 10425.5; DB 15; Length 2066;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 2044; Conservative 2; Mismatches 5; Indels 1; Gaps 1;

1 MLKPKYGRVPLDGAEPILASRSRLNLFPOGKPPMTQOQMSPLSREGILALFVLFE 60
1 MLKPKYGRVPLDGAEPILASRSRLNLFPOGKPPMTQOQMSPLSREGILALFVLFE 60
61 ECSPALMKTIVHSNFFVRKYSDDTIAELQELQPSAKDFEVSILVCGHFAEYQVVRKATG 120
61 ECSPALMKTIVHSNFFVRKYSDDTIAELQELQPSAKDFEVSILVCGHFAEYQVVRKATG 120
121 DIYAMKWKKKALIAQOVSPFEEERNILSRSTSPWIPOLQYAFQDNHLYLVMEYOPGG 180
121 DIYAMKWKKKALIAQOVSPFEEERNILSRSTSPWIPOLQYAFQDNHLYLVMEYOPGG 180
121 DIYAMKWKKKALIAQOVSPFEEERNILSRSTSPWIPOLQYAFQDNHLYLVMEYOPGG 180
181 DLISLNRVEDQDENLLOFYLAELILAVHSVHLMGVHVRDIRENLLVDRTHIKLVDF 240
181 DLISLNRVEDQDENLLOFYLAELILAVHSVHLMGVHVRDIRENLLVDRTHIKLVDF 240
181 DLISLNRVEDQDENLLOFYLAELILAVHSVHLMGVHVRDIRENLLVDRTHIKLVDF 240
241 GSAKONSNMKNVNAKLPIGTPTYAPREVLTVMNGDGKTYGLDCDMSVGVIAVEMTYGR 300
241 GSAKONSNMKNVNAKLPIGTPTYAPREVLTVMNGDGKTYGLDCDMSVGVIAVEMTYGR 300
241 GSAKONSNMKNVNAKLPIGTPTYAPREVLTVMNGDGKTYGLDCDMSVGVIAVEMTYGR 300
301 SPFAEGTSARTFNINMFORFLKFPDDPKVSSDFLILQSLICGOKERLKFEGICHPFF 360
301 SPFAEGTSARTFNINMFORFLKFPDDPKVSSDFLILQSLICGOKERLKFEGICHPFF 360
300 SPFAEGTSARTFNINMFORFLKFPDDPKVSSDFLILQSLICGOKERLKFEGICHPFF 359
361 SKIDMNNIRNSPPFVTVTLKSDDDTSNFDPEKSNWSSPCOLSPSGFGEELPFVGF 420
361 SKIDMNNIRNSPPFVTVTLKSDDDTSNFDPEKSNWSSPCOLSPSGFGEELPFVGF 420
360 SKIDMNNIRNSPPFVTVTLKSDDDTSNFDPEKSNWSSPCOLSPSGFGEELPFVGF 419
421 YSKALGILGRSESVGLDSPAKTSMKKLLIKSKELQDSODKCHMEQEMTLHRVS 480
421 YSKALGILGRSESVGLDSPAKTSMKKLLIKSKELQDSODKCHMEQEMTLHRVS 480
420 YSKALGILGRSESVGLDSPAKTSMKKLLIKSKELQDSODKCHMEQEMTLHRVS 479
481 EYAVAVLSQKEVELKASTQSLLEQDILATYITTECSILKRSLEQARMVEVSQEDDKALQLH 540
481 EYAVAVLSQKEVELKASTQSLLEQDILATYITTECSILKRSLEQARMVEVSQEDDKALQLH 540
480 EYAVAVLSQKEVELKASTQSLLEQDILATYITTECSILKRSLEQARMVEVSQEDDKALQLH 539
541 DIBBQSKLOEIKQEQVQAVVEEMRLMNOLEBDLVARRRSDLYESELLESRLAAEFK 600
541 DIBBQSKLOEIKQEQVQAVVEEMRLMNOLEBDLVARRRSDLYESELLESRLAAEFK 600
540 DIBBQSKLOEIKQEQVQAVVEEMRLMNOLEBDLVARRRSDLYESELLESRLAAEFK 599

601 RKATCQHKLLKAKDQKPEVGEYAVLCEKINAEOQLKIOELQELKAVKASTTEATELLQ 660
600 RKATCQHKLLKAKDQKPEVGEYAVLCEKINAEOQLKIOELQELKAVKASTTEATELLQ 659
661 NIFQAKERARELEKIQNREDSSEGIKKLVAEERHSHLENVKSLETERENRLLKDD 720
660 NIFQAKERARELEKIQNREDSSEGIKKLVAEERHSHLENVKSLETERENRLLKDD 719
721 IQTSQOIQMADKILIELEKHEAQSVAQHLVEHLKQKQHYEKKIKVLNDQIKDDLD 780
720 IQTSQOIQMADKILIELEKHEAQSVAQHLVEHLKQKQHYEKKIKVLNDQIKDDLD 779
781 KETLENNMQHEEBAHEKGIKISEQYAMINADSKIRSLFORIVELSEANKLAANSFLT 840
780 KETLENNMQHEEBAHEKGIKISEQYAMINADSKIRSLFORIVELSEANKLAANSFLT 839
841 QRNMKAQOEEMISLROOKFLETQAKLEAONKLEEOLEKISHOHSQNRLLLETRL 900
840 QRNMKAQOEEMISLROOKFLETQAKLEAONKLEEOLEKISHOHSQNRLLLETRL 899
901 REVSLHEEQKLELQTELQSLQERESQTLALQAAALAESQLOAKTELEETTAEA 960
900 REVSLHEEQKLELQTELQSLQERESQTLALQAAALAESQLOAKTELEETTAEA 959
961 EEEIOALTARDEIQKFDALRNSCTVITDLSEBOLNLTEDNAELNNQNFYSKQIDEAS 1020
960 EEEIOALTARDEIQKFDALRNSCTVITDLSEBOLNLTEDNAELNNQNFYSKQIDEAS 1019
1021 GANDEIVOLRSEVDHRRREITREEMQTSOKQMEALKTCTMLEBOWMDLEALNLELE 1080
1020 GANDEIVOLRSEVDHRRREITREEMQTSOKQMEALKTCTMLEBOWMDLEALNLELE 1079
1081 KERQWEMRVSVLGDEKSOFECEVRELORMIDTEKQSRABADQRTESRQVVELAVENKX 1140
1080 KERQWEMRVSVLGDEKSOFECEVRELORMIDTEKQSRABADQRTESRQVVELAVENKX 1139
1141 EILALQALKEQKAKESLSDKINDLEKKHAMLENNARSIQQLTEFBREIKORLSEQAK 1200
1140 EILALQALKEQKAKESLSDKINDLEKKHAMLENNARSIQQLTEFBREIKORLSEQAK 1199
1201 LQOQMDLOKNHIFRLTQGLQEALDRADLLKTESDLEQULENTOVLYSHKVMWBGTSIQ 1260
1200 LQOQMDLOKNHIFRLTQGLQEALDRADLLKTESDLEQULENTOVLYSHKVMWBGTSIQ 1259
1261 QTKLIDFLQAKMDQPAKKKKVPLQVNEMLKLALEKAKACALEBEALQKTRIELBSAREEA 1320
1260 QTKLIDFLQAKMDQPAKKKKVPLQVNEMLKLALEKAKACALEBEALQKTRIELBSAREEA 1319
1321 AHRKATDHPHSTPATARQOIAMSAIVRSPEHQPSAMSLAPSSSRKKSSTPEEFSRRL 1380
1320 AHRKATDHPHSTPATARQOIAMSAIVRSPEHQPSAMSLAPSSSRKKSSTPEEFSRRL 1379
1381 KERMHNNIPHRFVNGVNMRAATKCAVCLDVHFRQASKCLECQWCHPKSCSTLPLATCG 1440
1380 KERMHNNIPHRFVNGVNMRAATKCAVCLDVHFRQASKCLECQWCHPKSCSTLPLATCG 1439
1441 PAEYATHTFAFCRDQNSPGLQTKPESSSLHLEGMMKYPNNKRKQOGQMDRKYIVLEGS 1500
1440 PAEYATHTFAFCRDQNSPGLQTKPESSSLHLEGMMKYPNNKRKQOGQMDRKYIVLEGS 1499
1501 KVLIDYNEAREAGQRFVEEPELCLPGDVSIGAVASSELANTAKADVPIYILMESHPHT 1560
1500 KVLIDYNEAREAGQRFVEEPELCLPGDVSIGAVASSELANTAKADVPIYILMESHPHT 1559
1561 TCMPSGRTYILASFPDQKQWWTALSVVAGGVSEKKAADKLLGNSLLKLBGGDRLD 1620
1560 TCMPSGRTYILASFPDQKQWWTALSVVAGGVSEKKAADKLLGNSLLKLBGGDRLD 1619
1621 MCTLPSPSOVVLVGEEGLYALNVLKNSLTHVPGIGAVFOYIYIKOLEKLMIAEGERA 1680
1620 MCTLPSPSOVVLVGEEGLYALNVLKNSLTHVPGIGAVFOYIYIKOLEKLMIAEGERA 1679

QY 1681 LCLVDVKKVKSLOSLAOSHLPADPDISPNI FEAVKCGHLPAGKINGLCTICAMPSKVITL 1740
Db 1680 LCLVDVKKVKSLOSLAOSHLPADPDISPNI FEAVKCGHLPAGKINGLCTICAMPSKVITL 1739
QY 1741 RYNNENLSKYCIKRIELETSEPCSHFTNYSLIGTNKFEYIDMKOYTLLEPLDKNDHSLA 1800
Db 1740 RYNNENLSKYCIKRIELETSEPCSHFTNYSLIGTNKFEYIDMKOYTLLEPLDKNDHSLA 1739
QY 1801 PAVFAASNSFPVSIYOVNSAGOREEYLLCFHEFGVPVDSYGRSRTDCLKMSRLPLAFA 1860
Db 1800 PAVFAASNSFPVSIYOVNSAGOREEYLLCFHEFGVPVDSYGRSRTDCLKMSRLPLAFA 1859
QY 1861 YREBYLFTVTHNSLEVEIEIOARSSAGTPARAYLIDIPNRYLGPALISSAATILASYYDKL 1920
Db 1860 YREBYLFTVTHNSLEVEIEIOARSSAGTPARAYLIDIPNRYLGPALISSAATILASYYDKL 1919
QY 1921 RVICCKGNLVKESGTEHHRGPSTSRSSPNKGPPTVNEHTTKRAVASSPAPPEGSHPREP 1980
Db 1920 RVICCKGNLVKESGTEHHRGPSTSRSSPNKGPPTVNEHTTKRAVASSPAPPEGSHPREP 1979
QY 1981 STEPHREGETELRDSPGRPLERKSPGRILSTRERSPARLFEDESSRGLPAGAVRT 2040
Db 1980 STEPHREGETELRDSPGRPLERKSPGRILSTRERSPARLFEDESSRGLPAGAVRT 2039
QY 2041 PLSQVANKVMDOS 2052
Db 2040 PLSQVANKVMDOS 2051

RESULT 8

US-10-017-216-4
; Sequence 4, Application US/10017216
; Publication No. US20020160483A1
; GENERAL INFORMATION:
; APPLICANT: KAPLELLER-LIBERMAN, Rosana
; TITLE OF INVENTION: 13245, A No. US20020160483A1 Human Myotonic Dystrophy Type Proc
; TITLE OF INVENTION: Kinase and Uses Therefor
; FILE REFERENCE: 10147-5701
; CURRENT APPLICATION NUMBER: US/10/017, 216
; CURRENT FILING DATE: 2001-10-23
; PRIOR APPLICATION NUMBER: US 60/242,429
; PRIOR FILING DATE: 2000-10-23
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 4
; LENGTH: 2055
; TYPE: PRN
; ORGANISM: Mus musculus
US-10-017-216-4

Query Match 96.3%; Score 10109.5; DB 13; Length 2055;
Best Local Similarity 96.0%; Pred. No. 0;
Matches 1974; Conservative 40; Mismatches 39; Indels 3; Gaps 2;
QY 1 MLKFTYGARNPLDGAAPLASRASRLNLFQGRPPTTQOQMSPLSREGILDLFVLPFE 60
Db 1 MLKFTYGVNRUPPEASASEPIASRASRLNLFQGRPPTTQOQMSPLSREGILDLFVLPFE 60
QY 61 ECSPALAMKIKHVNPNFRKYSDDTIAELQELOPSADKDEEVNSLVCCGFAEVQVVRKATG 120
Db 61 ECSPALAMKIKHVNPNFRKYSDDTIAELQELOPSADKDEEVNSLVCCGFAEVQVVRKATG 120
QY 121 DIYAMKVMKKKALLAOQVSPFEERNTLSRSTSPMIPOLQYAFQDKNHLVLMVEYOPGG 180
Db 121 DIYAMKVMKKKALLAOQVSPFEERNTLSRSTSPMIPOLQYAFQDKNHLVLMVEYOPGG 180
QY 181 DLISLNRREYDQDENTLIQFYLAELILAVSHVMGVVHRDIKENTILVDRGTGHIKLVDF 240
Db 181 DLISLNRREYDQDENTLIQFYLAELILAVSHVMGVVHRDIKENTILVDRGTGHIKLVDF 240
QY 241 GSAAKMNSNMVNAKLPICGTPDYMAPEVLTVMEDRRGTGLDCDMSVGVAVAYEMVYK 300
Db 241 GSAAKMNSNMVNAKLPICGTPDYMAPEVLTVMEDRRGTGLDCDMSVGVAVAYEMVYK 299

QY 301 SPFAEGTSARTENNINMNFQRFLEKPPDDPKVSSDFLDLISLLCQOKERLEKEGLCCHPEF 360
Db 300 TPFTEGTSARTENNINMNFQRFLEKPPDDPKVSSDFLDLISLLCQOKERLEKEGLCCHPEF 359
QY 361 SKIMNNINRSPPFVTLSDDDTNSNPFDEKNSWVSSPCQSLPSPGSGEELPFVGF 420
Db 360 ARTMNNINRSPPFVTLSDDDTNSNPFDEKNSWVSSPCQSLPSPGSGEELPFVGF 419
QY 421 YSKALGILLGSESVSGDSPAKTSMEKLLISKELODQDQCHMKQEMTRLHRVVS 480
Db 420 YSKALGILLGSESVSGDSPAKTSMEKLLISKELODQDQCHMKQEMTRLHRVVS 479
QY 481 EFAVALSQKEVELKASSTQSLQSLDATTYITTECSSLKRSLEQARMEVSQEDDKALQLH 540
Db 480 EFAVALSQKEVELKASSTQSLQSLDATTYITTECSSLKRSLEQARMEVSQEDDKALQLH 539
QY 541 DIREOSRKLQEIKEQEOYAOVEEMRLMNOLEEDLVASARRSDLYESELRESRLAAEFK 600
Db 540 DIREOSRKLQEIKEQEOYAOVEEMRLMNOLEEDLVASARRSDLYESELRESRLAAEFK 599
QY 601 RKATECOHKLKAKDQCKPEVEGYAKLEKINAEOQLKIQELQELKRAVASTATELLQ 660
Db 600 RKANECHKLKAKDQCKPEVEGYAKLEKINAEOQLKIQELQELKRAVASTATELLQ 659
QY 661 NIROAKERAERLEKLNREDSSBGRKCLVEABERHSHLENKYRLTETMERENRLKD 720
Db 660 NIROAKERAERLEKLNREDSSBGRKCLVEABERHSHLENKYRLTETMERENRLKD 719
QY 721 IQTSQOIQOQADKILLEEKHRAOVSAGHLEVHLQKQOHEBEKIKVLDNQIKCDLAD 780
Db 720 IQTSQOIQOQADKILLEEKHRAOVSAGHLEVHLQKQOHEBEKIKVLDNQIKCDLAD 779
QY 781 KETLENNMOHREBEAHEKGLISQKAMINAMDSKISLBORIYELSEANKLAANSLSFT 840
Db 780 KESLENNMOHREBEAHEKGLISQKAMINAMDSKISLBORIYELSEANKLAANSLSFT 839
QY 841 QNNKAQEMISSELROQFYLETOAGCLEAONRKLBEQLEKISHQDSDKNRLLELETRL 900
Db 840 QNNKAQEMISSELROQFYLETOAGCLEAONRKLBEQLEKISHQDSDKNRLLELETRL 899
QY 901 REVSLHEBEQLEKROULTLEQSLQERESQTLQAARALBSQLRQAKTELEETTAEA 960
Db 900 REVSLHEBEQLEKROULTLEQSLQERESQTLQAARALBSQLRQAKTELEETTAEA 959
QY 961 EBEIOALTARDEIQRFDALRNSCTVITDLEBQNLQTEDNAELNNQNFPLSQOLBAS 1020
Db 960 EBEIOALTARDEIQRFDALRNSCTVITDLEBQNLQTEDNAELNNQNFPLSQOLBAS 1019
QY 1021 GANDEIVQLRSEVDHLREITEREMOULTSQKQTEALKTCTMLEBQVMDLEALNDELE 1080
Db 1020 GANDEIVQLRSEVDHLREITEREMOULTSQKQTEALKTCTMLEBQVMDLEALNDELE 1079
QY 1081 KEROMEAMRSVLDGKESQFEQVREILOPMLDTEKQSPARADORTESROVVELAVKCHKA 1140
Db 1080 KEROMEAMRSVLDGKESQFEQVREILOPMLDTEKQSPARADORTESROVVELAVKCHKA 1139
QY 1141 EILALQOLKBEQKKAESLSDKLNDEKKNAMLENNARSLOQLKETERELKORLLEBOAK 1200
Db 1140 EILALQOLKBEQKKAESLSDKLNDEKKNAMLENNARSLOQLKETERELKORLLEBOAK 1199
QY 1201 LQOQMDLOKNHIFLTLQLOEALDRADLTKTERSDELYOLENTIOVLSHEKVKMEGTISQ 1260
Db 1200 LQOQMDLOKNHIFLTLQLOEALDRADLTKTERSDELYOLENTIOVLSHEKVKMEGTISQ 1259
QY 1261 QTKLIDFLQAKMDQPAKKKVPLOYNELKALEKAKACABLEALOKTRIELRSARBEA 1320
Db 1260 QTKLIDFLQAKMDQPAKKKVPLOYNELKALEKAKACABLEALOKTRIELRSARBEA 1319
QY 1321 AHRKATDHPSPPTATROQIAMSATVRSPEHOPASMSLLAPSSRKESSTPEEFSRL 1380
Db 1320 AHRKATDHPSPPTATROQIAMSATVRSPEHOPASMSLLAPSSRKESSTPEEFSRL 1379

QY 1381 KERMHNIPHRFVNLNRAITKCAVCLDTHFGQASKCIECOVMCHPKSTCIPATCGL 1440
Db 1380 KERMHNIPHRFVNLNRAITKCAVCLDTHFGQASKCIECOVMCHPKSTCIPATCGL 1439
QY 1441 PABYATHTFAFCDKNNSPGLQTKPSSSLHLEGMMKVPRNNRGGQGDRLKTVLEGS 1500
Db 1440 PABYATHTFAFCDKNNSPGLQTKPSSSLHLEGMMKVPRNNRGGQGDRLKTVLEGS 1499
QY 1501 KVLIDNEAREAGORPVEEFELCLPDGDVSIHGAVGASELANAKADVPIYLKXNESHPT 1560
Db 1500 KVLIDNEAREAGORPVEEFELCLPDGDVSIHGAVGASELANAKADVPIYLKXNESHPT 1559
QY 1561 TCMPTGRTLYLLASFPDKXWMTALBSVAVAGGRVSRKABADAKLGNLSLKLEGGDRLD 1620
Db 1560 TCMPTGRTLYLLASFPDKXWMTALBSVAVAGGRVSRKABADAKLGNLSLKLEGGDRLD 1619
QY 1621 MNCITLPSDQVLYVGTSEGLYALNVNLKNSLTHVPGIGAVFOIYIILKLEKMLMAGEERA 1680
Db 1620 MNCITLPSDQVLYVGTSEGLYALNVNLKNSLTHVPGIGAVFOIYIILKLEKMLMAGEERA 1679
QY 1681 LCLVDYKVKQSLAQSHLPAQDISPNIPEAVKGCGLFGAGKLENGLCICAMPSKVYL 1740
Db 1680 LCLVDYKVKQSLAQSHLPAQDISPNIPEAVKGCGLFGAGKLENGLCICAMPSKVYL 1739
QY 1741 RYNNENSKYICRKEIETSEPCSIHFTNYSILIGTNKPYEIDMKQYTLLEFPLDKNDHSLA 1800
Db 1740 RYNNENSKYICRKEIETSEPCSIHFTNYSILIGTNKPYEIDMKQYTLLEFPLDKNDHSLA 1799
QY 1801 PAVPAASNSPPVSIYOVNAGOREEYLLCFHERGVVDVSGRRSRDDLKMSLPLAFA 1860
Db 1800 PAVPAASNSPPVSIYOVNAGOREEYLLCFHERGVVDVSGRRSRDDLKMSLPLAFA 1859
QY 1861 YREBYLVTHTFNSLVELEIQAASSAGTAPAYLDIPNRYLGPALSSGATYLAASYODKL 1920
Db 1860 YREBYLVTHTFNSLVELEIQAASSAGTAPAYLDIPNRYLGPALSSGATYLAASYODKL 1919
QY 1921 RVICCKGNLVKSGTEHHRGPTSRSSPNKRGPTTNEHTTKRVASSPAPPEGSHPREP 1980
Db 1920 RVICCKGNLVKSGTEHHRGPTSRSSPNKRGPTTNEHTTKRVASSPAPPEGSHPREP 1979
QY 1981 STPHRY--RGRGTRLRDCKSPGRPLERKSPGRILSRRRSPRRLLEDSRGLPLGAV 2038
Db 1980 STPHRYRDRGRTELRDCKSPGRPLERKSPGRMLSTRRERSPEGLLEDSRGLPLGAV 2039
QY 2039 RTPLSQVNAKYWDOSV 2054
Db 2040 RTPLSQVNAKYWDOSV 2055

RESULT 9
US-10-017-216-2
; Sequence 2, Application US/10017216
; Publication No. US20020160483A1
; GENERAL INFORMATION:
; APPLICANT: KABELLER-LIBERMANN, Rosana
; TITLE OF INVENTION: 13245, A No. US20020160483A1el Human Myoclonic Dystrophy Type Prot
; FILE REFERENCE: 10147-57U1
; CURRENT APPLICATION NUMBER: US/10/017, 216
; CURRENT FILING DATE: 2001-10-23
; PRIOR APPLICATION NUMBER: US 60/242, 429
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 2053
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-017-216-2

Query Match 95.5%; Score 10022.5; DB 13; Length 2053;
Best Local Similarity 95.7%; Pred. No. 0;
Matches 1979; Conservative 3; Mismatches 5; Indels 81; Gaps 3;

QY 1 MKFKYGAENPLDAGAEPIASPASLNLFPQCKPPMTQOQNSPLSRBSGLDALFTVLE 60
Db 1 MKFKYGAENPLDAGAEPIASPASLNLFPQCKPPMTQOQNSPLSRBSGLDALFTVLE 60
QY 61 ECGOPALMKIKHNSNVKRYSDTIAELOPQAKDPEVRSILVGCCHFAEVOVREKATG 120
Db 61 ECGOPALMKIKHNSNVKRYSDTIAELOPQAKDPEVRSILVGCCHFAEVOVREKATG 120
QY 121 DIYAMVKMKKALLAAEOVSFFEEERNILSRSTSPMIPOLQVAFODKNMLYLVMEYQPG 180
Db 121 DIYAMVKMKKALLAAEOVSFFEEERNILSRSTSPMIPOLQVAFODKNMLYLVMEYQPG 180
QY 181 DLISLNRVEDQDENLQFYLAELILAVSHVLMGYVRDIPENILVDRTHGILKLVDF 240
Db 181 DLISLNRVEDQDENLQFYLAELILAVSHVLMGYVRDIPENILVDRTHGILKLVDF 240
QY 241 GSAAKNNSNMVNAKPIGTPDYMAPEVLTVMNGDGKGTGLCDMMSVGVIAEMTYGR 300
Db 241 GSAAKNNSNMVNAKPIGTPDYMAPEVLTVMNGDGKGTGLCDMMSVGVIAEMTYGR 300
QY 301 SPFAEGTSARTFNNINMFORFLKFPDDPKVSDPLDILQSLCGQERLKEBGLCHPFF 360
Db 301 SPFAEGTSARTFNNINMFORFLKFPDDPKVSDPLDILQSLCGQERLKEBGLCHPFF 360
QY 361 SKIDMNNIRNSPPPVPTLKSDDDTSNFDEPEKNSVSSPCQLSPSGSGEBLPHYGFS 420
Db 361 SKIDMNNIRNSPPPVPTLKSDDDTSNFDEPEKNSVSSPCQLSPSGSGEBLPHYGFS 420
QY 421 YSKALGILGRSEVSGLSPATSSMEKKLILKSKELODSQCKHMEDEMRLLHRYS 480
Db 421 YSKALGILGRSEVSGLSPATSSMEKKLILKSKELODSQCKHMEDEMRLLHRYS 480
QY 481 EYAVAVLSQKEVELKASETORSLLEODLATYITECSSLKSLEQARNEVSGEDKALQLH 540
Db 481 EYAVAVLSQKEVELKASETORSLLEODLATYITECSSLKSLEQARNEVSGEDKALQLH 540
QY 541 DIREQSRKLOEIKOEYQAOVEBRLMMNQLBEDLVASARRSULYSELRESLAAEFK 600
Db 541 DIREQSRKLOEIKOEYQAOVEBRLMMNQLBEDLVASARRSULYSELRESLAAEFK 600
QY 601 RKAETCOHKLKAKDQKPEVGEYALKETKMAEQOLKIOLOKRLKAVASTEATVLLQ 660
Db 601 RKAETCOHKLKAKDQKPEVGEYALKETKMAEQOLKIOLOKRLKAVASTEATVLLQ 660
QY 661 NIRAQKARERLELEKQNEREDSSEGIKKLVLEAEBRRHSLENKVKLETERRENRKOD 720
Db 661 NIRAQKARERLELEKQNEREDSSEGIKKLVLEAEBRRHSLENKVKLETERRENRKOD 720
QY 721 IOTKSOQIQOMADKILLEBKREAOVSAQHLEVHLKQKQHYEYKIKVLNDQIKODLAD 780
Db 721 IOTKSOQIQOMADKILLEBKREAOVSAQHLEVHLKQKQHYEYKIKVLNDQIKODLAD 780
QY 781 KETLENMMQHRHEAEAEKGLISEQKAMINAMSKIRSEQRIVELSEAKKLANSLSFT 840
Db 781 KETLENMMQHRHEAEAEKGLISEQKAMINAMSKIRSEQRIVELSEAKKLANSLSFT 840
QY 841 QRMKAQOEKISELROOKFYLETQAGKLEBAONKLEEOLEKISHQDSDNRLLLEETRL 900
Db 841 QRMKAQOEKISELROOKFYLETQAGKLEBAONKLEEOLEKISHQDSDNRLLLEETRL 900
QY 901 REVSLHEBQKELKQUTELQSLQERSQUTALQAPRAALQSOLRQAKTELEETTAEA 960
Db 901 REVSLHEBQKELKQUTELQSLQERSQUTALQAPRAALQSOLRQAKTELEETTAEA 960
QY 961 EBEIOLALTARDEIQKFPDALRNSCTVITDLSEQLNQLTEDNAELNNQNFYSKQDEAS 1020
Db 961 EBEIOLALTARDEIQKFPDALRNSCTVITDLSEQLNQLTEDNAELNNQNFYSKQDEAS 1004
QY 1021 GANDEIVOLRSEVDHLRREITEREMQUTSQKOTMEALKTCTYALSEQVNDLEALNDELLE 1080
Db 1005 GANDEIVOLRSEVDHLRREITEREMQUTSQKOTMEALKTCTYALSEQVNDLEALNDELLE 1064

1081 KEROMEARSVLGDKEKQFEGCRVRELOMLDTEKOSRARADORTESHQVVELAKHEKA 1140
1065 KEROMEANRSVLGBEKSQFECRVELQMLDTEKOSRARADORTESHQVVELAKHEKA 1124
1141 EIALAQAALKEQKLAESLSDKLANDLEKHAMLEMANASLOOKLETREBELQORLLEQAK 1200
1125 EIALAQAALKEQKLAESLSDKLANDLEKHAMLEMANASLOOKLETREBELQORLLEQAK 1184
1201 LQOQMDLOKNIHIFLITQGLQALDRADLLKTERSDLEYQLENIOVLVSEHEKVEGTISQ 1260
1185 LQOQMDLOKNIHIFLITQGLQALDRADLLKTERSDLEYQLENIOVLVSEHEKVEGTISQ 1244
1261 QTKLIDFLQAKMDQPAKKKK-----VPLQVNEKLALKEKARCALEBEA 1305
1245 QTKLIDFLQAKMDQPAKKKKGLFSRRKEDPALPTQVPLQVNEKLALKEKARCALEBEA 1304
1306 LQKTRIELRSAREBAARHAKATDHPHSTPATAROOIAMSALVRSPEHOPSAMSLIAPSS 1365
1305 LQKTRIELRSAREBAARHAKATDHPHSTPATAROOIAMSALVRSPEHOPSAMSLIAPSS 1364
1366 RRKESSTPEEFSRRLKERMHNIIPHRFVNGINMRATKCAVCLDTVHFGROASKLECOVM 1425
1365 RRKESSTPEEFSRRLKERMHNIIPHRFVNGINMRATKCAVCLDTVHFGROASKLECOVM 1424
1426 CHPKCSTCLPATCGLPAYATHTFTEAFGRDKMNSPGLQTKPESSSLHEGMKTPRNKR 1485
1425 CHPKCSTCLPATCGLPAYATHTFTEAFGRDKMNSPGLQTKPESSSLHEGMKTPRNKR 1484
1486 GQOQMDRKYIYLEGSKVLIYNEAREAGORVPEEBELCLPGDVS IHGAVASLANAK 1545
1485 GQOQMDRKYIYLEGSKVLIYNEAREAGORVPEEBELCLPGDVS IHGAVASLANAK 1544
1546 ADVPEYILKMHESHPTTCMGRITLYLLAFSPFDKORWYALLESVVAAGVSRSEKADAKL 1605
1545 A-----EKEMADAKL 1554
1606 LGNSLLKLEGGDRDLDMNCTLPFSQOVVLVGTBEGLYALNVLKNSLTHVPGIAGVFOIYI 1665
1555 LGNSLLKLEGGDRDLDMNCTLPFSQOVVLVGTBEGLYALNVLKNSLTHVPGIAGVFOIYI 1614
1666 KDLKELMIMAGEBALCLVDKVKYKOSLAOSHLPQDPISNITEAVAGCHLFGAGKIEN 1725
1615 KDLKELMIMAGEBALCLVDKVKYKOSLAOSHLPQDPISNITEAVAGCHLFGAGKIEN 1674
1726 GLCICAMPKSVLLIARVNEMLSKYCIKKELETSEPCSIHFTNYSILIGTKKFEIEMQ 1785
1675 GLCICAMPKSVLLIARVNEMLSKYCIKKELETSEPCSIHFTNYSILIGTKKFEIEMQ 1734
1786 YTLBEFLDKXNDHSLAPVFAASSNSPVSIQVNSAQREBYLLCFHEFGVFDVSGRS 1845
1735 YTLBEFLDKXNDHSLAPVFAASSNSPVSIQVNSAQREBYLLCFHEFGVFDVSGRS 1794
1846 RTDLDKKSRLPLAFAYREPLFTVHNSLEYIELOASSAGTPARAYLIDIPNRYLGPAL 1905
1795 RTDLDKKSRLPLAFAYREPLFTVHNSLEYIELOASSAGTPARAYLIDIPNRYLGPAL 1854
1906 SSGAIYLAASYODKLARIYICCKGNLVKSGTEHHRGPTSSRSPKRGPPYTNIEITRVA 1965
1855 SSGAIYLAASYODKLARIYICCKGNLVKSGTEHHRGPTSSRSPKRGPPYTNIEITRVA 1914
1966 SSPAPPEGSHPRPSTPHRYREGRTLRDSDGRLERKSPGRLILSTRERSPARLP 2025
1915 SSPAPPEGSHPRPSTPHRYREGRTLRDSDGRLERKSPGRLILSTRERSPARLP 1974
2026 EDSRRGLPLAGAVRTPILSOVNVKWDSS 2053
1975 EDSRRGLPLAGAVRTPILSOVNVKWDSS 2002

RESULT 10
US-10-325-430-12
; Sequence 12, Application US/10325430
; Publication No. US20030153525A1

GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc
; APPLICANT: Silos-Santiago, Immaculada
; APPLICANT: Rosenfeld, Julie Beth
; TITLE OF INVENTION: METHODS AND COMPOSITIONS IN TREATING
; TITLE OF INVENTION: PAIN AND PAINFUL DISORDERS USING 1465, 1587, 2146, 2207,
; TITLE OF INVENTION: 32838,336 and 52908
; FILE REFERENCE: MP101-294P1RNM
; CURRENT APPLICATION NUMBER: US/10/325,430
; CURRENT FILING DATE: 2002-12-19
; PRIOR APPLICATION NUMBER: US 60/341,953
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 2053
; TYPE: PR
; ORGANISM: Homo Sapiens
US-10-325-430-12
Query Match 95.5%; Score 10022.5; DB 14; Length 2053;
Best Local Similarity 95.7%; Pred. No. 0;
Matches 1979; Conservative 3; Mismatches 5; Indels 81; Gaps 3;
QY 1 MLKRYGARNPLDGAAPLASRASRLNLFQKRPMTQOQMSPLSREGILDLFVLFE 60
DB 1 MLKRYGARNPLDGAAPLASRASRLNLFQKRPMTQOQMSPLSREGILDLFVLFE 60
QY 61 ECGSPALMKIKHVSNPFRKYSDDTIAELOOPSADKEVRELVCGGHFAEVQVVRKATG 120
DB 61 ECGSPALMKIKHVSNPFRKYSDDTIAELOOPSADKEVRELVCGGHFAEVQVVRKATG 120
QY 121 DIYMKYMKKKALLAQOVSFFEEERNILSRSTPMIPOLQYARODKNHLYLMEYOPCG 180
DB 121 DIYMKYMKKKALLAQOVSFFEEERNILSRSTPMIPOLQYARODKNHLYLMEYOPCG 180
QY 121 DIYMKYMKKKALLAQOVSFFEEERNILSRSTPMIPOLQYARODKNHLYLMEYOPCG 180
DB 121 DIYMKYMKKKALLAQOVSFFEEERNILSRSTPMIPOLQYARODKNHLYLMEYOPCG 180
QY 181 DLISLNRVEDQDLENLIFOYLAELILAVHSVHLMGVHRDIKENLIVDRTHIKLVDF 240
DB 181 DLISLNRVEDQDLENLIFOYLAELILAVHSVHLMGVHRDIKENLIVDRTHIKLVDF 240
QY 241 GSAAKMNSNMVNAKLIGTPDYNAPEVLTYMNDGKGTGLDDDMNSVGIAYEMTYGR 300
DB 241 GSAAKMNSNMVNAKLIGTPDYNAPEVLTYMNDGKGTGLDDDMNSVGIAYEMTYGR 300
QY 241 GSAAKMNSNMVNAKLIGTPDYNAPEVLTYMNDGKGTGLDDDMNSVGIAYEMTYGR 300
DB 241 GSAAKMNSNMVNAKLIGTPDYNAPEVLTYMNDGKGTGLDDDMNSVGIAYEMTYGR 300
QY 301 SPFAEGTSATFNINMFORFLKEPDDPKVSDPLDIQSILCGQKELKFEGLCCHPFF 360
DB 301 SPFAEGTSATFNINMFORFLKEPDDPKVSDPLDIQSILCGQKELKFEGLCCHPFF 360
QY 361 SKIDMNNIRNSPPEFVPTLKSDDDTSNFDEBKNSWVSSPCQLSPSGFSGEBLPHYGFS 420
DB 361 SKIDMNNIRNSPPEFVPTLKSDDDTSNFDEBKNSWVSSPCQLSPSGFSGEBLPHYGFS 420
QY 421 YSKALGILGSESVSGLDSPAKTSMEKKLLYSKELQSDQCHKMEQMTLHRRVS 480
DB 421 YSKALGILGSESVSGLDSPAKTSMEKKLLYSKELQSDQCHKMEQMTLHRRVS 480
QY 481 EVEAVLSQKEVELKASRTOSLLEODLATYITTECSSIKRSLEQARMVSOQDDKALQILH 540
DB 481 EVEAVLSQKEVELKASRTOSLLEODLATYITTECSSIKRSLEQARMVSOQDDKALQILH 540
QY 541 DIBQSRKLOEIKOEYOQAVVEEMRLMNNQLEEDLVASARRSDLYESELRESRLAAEFK 600
DB 541 DIBQSRKLOEIKOEYOQAVVEEMRLMNNQLEEDLVASARRSDLYESELRESRLAAEFK 600
QY 601 RKAITECHKLKAKDOQKPEVGEYAKLEKINAEQOLKIOELOEKLAVASTATELLO 660
DB 601 RKAITECHKLKAKDOQKPEVGEYAKLEKINAEQOLKIOELOEKLAVASTATELLO 660
QY 661 NIKOAKERARELEKLONREDSSEGIKRLVBAEERHSLJENKYKRLJETERRENRRLKDD 720
DB 661 NIKOAKERARELEKLONREDSSEGIKRLVBAEERHSLJENKYKRLJETERRENRRLKDD 720
QY 721 IQTSQOQIQOMADKILEBEKHREAOVSQHLBVLHKKOEHYEEKIKVLNIOIKDLAD 780
DB 721 IQTSQOQIQOMADKILEBEKHREAOVSQHLBVLHKKOEHYEEKIKVLNIOIKDLAD 780

Db 705 IQRTSQIQQMAKRIILEEKHREAOVSAOHLFEVHLQKQEHVEEKIKVLDNQKXDLAD. 764
Qy 781 KETLENNMOHSEEAHHEKGISLROKAMINAMDSKIRSLBORIYELSEANKLANSSLFT 840
Db 765 KETLENNMOHSEEAHHEKGISLROKAMINAMDSKIRSLBORIYELSEANKLANSSLFT 824
Qy 841 QRMNKAQEMISELROQKYLETOAGKLEAONRKLBOLEKISHQDSDKXRLLEETRL 900
Db 825 QRMNKAQEMISELROQKYLETOAGKLEAONRKLBOLEKISHQDSDKXRLLEETRL 884
Qy 901 REVSLEHEBOKLEIKROLTELQSLQERESQTLQARALBESQLOAKTELEETAA 960
Db 885 REVSLEHEBOKLEIKROLTELQSLQERESQTLQARALBESQLOAKTELEETAA 944
Qy 961 EEEIOLATLHARDEIQRKFADLRNSCTVITDLEBQNLQTEDNAELNNQNFYLSQOLBAS 1020
Db 945 EEEIOLATLHARDEIQRKFADLRNSCTVITDLEBQNLQTEDNAELNNQNFYLSQOLBAS 1004
Qy 1021 GANDEIVOLRSEVDHLRREITEREMOLTSOKQMEALKTCTMLEBOVMDLEALNDELLE 1080
Db 1005 GANDEIVOLRSEVDHLRREITEREMOLTSOKQMEALKTCTMLEBOVMDLEALNDELLE 1064
Qy 1081 KERQWEAMRSVLDGEKQFECRVELQPMILTEKQSPARADQRTTESQVVELAVKXKA 1140
Db 1065 KERQWEAMRSVLDGEKQFECRVELQPMILTEKQSPARADQRTTESQVVELAVKXKA 1124
Qy 1141 EITLALOALKBOKKASISLSDKNDLEKKAMLEMMNRSLOOKLETREIKORLLEBOAK 1200
Db 1125 EITLALOALKBOKKASISLSDKNDLEKKAMLEMMNRSLOOKLETREIKORLLEBOAK 1184
Qy 1201 LQOQMDLOKXNHFELTQGLQALDRADLKTERSDLEVOLENIQVLYSHEKVMEGTSIQ 1260
Db 1185 LQOQMDLOKXNHFELTQGLQALDRADLKTERSDLEVOLENIQVLYSHEKVMEGTSIQ 1244
Qy 1261 QTKLIDFLQAKMDQPAKKKK-----VPLQVNEKLALBKEKARCAELBEEA 1305
Db 1245 QTKLIDFLQAKMDQPAKKKKGLFSRRKEDPALPTQVPLQVNEKLALBKEKARCAELBEEA 1304
Qy 1306 LQKTRILRSAREEAHRAKATDHPSTPARAQAQIMSALVRSPELOPSMSSLAPSS 1365
Db 1305 LQKTRILRSAREEAHRAKATDHPSTPARAQAQIMSALVRSPELOPSMSSLAPSS 1364
Qy 1366 RRKESSTPEEFSRLKERMHNIPIHRFVGLNMATKCAVCLDPTVHGRQASKLECOVM 1425
Db 1365 RRKESSTPEEFSRLKERMHNIPIHRFVGLNMATKCAVCLDPTVHGRQASKLECOVM 1424
Qy 1426 CHPKSTCCLPATCGLPAYATHTFEAFCRDXKNSPGLQTKESPSSSLHEGMKVPRNNKR 1485
Db 1425 CHPKSTCCLPATCGLPAYATHTFEAFCRDXKNSPGLQTKESPSSSLHEGMKVPRNNKR 1484
Qy 1486 GQOQMDKRYIYLBESSKYLITYNEAREAGQRVVEFELCPGDVSHGAVASLANTAK 1545
Db 1485 GQOQMDKRYIYLBESSKYLITYNEAREAGQRVVEFELCPGDVSHGAVASLANTAK 1544
Qy 1546 ADVFYILKMEHSHPRITTCWGRRTLYILAPSPFDQKRWYALTESVAVAGRVSEKKAADAKL 1605
Db 1545 A-----EKHEADAKL 1554
Qy 1606 LGNSILKLBDGDRLDMMCTLPSPDQVVLVGTBEGLYALNVKNSLTHVPGI GAVFQIYI 1665
Db 1555 LGNSILKLBDGDRLDMMCTLPSPDQVVLVGTBEGLYALNVKNSLTHVPGI GAVFQIYI 1614
Qy 1666 KDLKELMIAGEBQALCLVDYKVKYQSLAQSHLPAQDPISNITREAVVGCFLFGAKIKEN 1725
Db 1615 KDLKELMIAGEBQALCLVDYKVKYQSLAQSHLPAQDPISNITREAVVGCFLFGAKIKEN 1674
Qy 1726 GLCTCAMPKSVVILIRYVENLSKYCIRKEIETSEPCSIHTFNYSILIGTKFYEIDMKQ 1785
Db 1675 GLCTCAMPKSVVILIRYVENLSKYCIRKEIETSEPCSIHTFNYSILIGTKFYEIDMKQ 1734
Qy 1786 YTLLEFLDKNDHSLAPAVFAASNSFPVSIQVNVNAGQREBYLLCFHEFGVFDVSGRRS 1845

Db 1735 YTLLEFLDKNDHSLAPAVFAASNSFPVSIQVNVNAGQREBYLLCFHEFGVFDVSGRRS 1794
Qy 1846 RTDDLKMSRLPLAFAYREPLYFTVHNSLEVEIEIOARSSAGTAPARAYLDIPNRYIGPAI 1905
Db 1795 RTDDLKMSRLPLAFAYREPLYFTVHNSLEVEIEIOARSSAGTAPARAYLDIPNRYIGPAI 1854
Qy 1906 SSGAIIYASSYQDKRLVYCKGNLYKESGTEHHRGSGTSSSNKGPPTVNEHITRVA 1965
Db 1855 SSGAIIYASSYQDKRLVYCKGNLYKESGTEHHRGSGTSSSNKGPPTVNEHITRVA 1914
Qy 1966 SSPAPPEGSHPREPSTPHRYREGRTLRDKSPGRPLEREKSPGRILSTRERSPARLF 2025
Db 1915 SSPAPPEGSHPREPSTPHRYREGRTLRDKSPGRPLEREKSPGRILSTRERSPARLF 1974
Qy 2026 EDSRGRPLPAGAVRTPLSQVNVWDQSS 2053
Db 1975 EDSRGRPLPAGAVRTPLSQVNVWDQSS 2002

RESULT 11
US-10-757-262-52
Sequence 52, Application US/10757262
Publication No. US20040197825A1
GENERAL INFORMATION:
APPLICANT: Karicneti, Venkateswarlu
APPLICANT: Eliasof, Scott D.
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING
UROLOGICAL DISORDERS USING 44390, 54181, 211, 5687, 884,
TITLE OF INVENTION: 1405, 636, 4421, 5410, 30905, 2045, 16405, 18560, 2047,
TITLE OF INVENTION: 33751, 52872, 14063, 20739, 32544, 43239, 44373, 51164,
TITLE OF INVENTION: 53010, 16852, 1587, 2207, 22245, 2387, 52908, 69112, 14990,
TITLE OF INVENTION: 18547, 115, 579, 15985, 15625, 760, 18603, 2395, 2554, 8675,
TITLE OF INVENTION: 32720, 4809, 14303, 16816, 17827, 32620, 5077, 619, 1423,
TITLE OF INVENTION: 2158, 8263, 15402, 16209, 16386, 21165, 30911, 41897, 1643,
TITLE OF INVENTION: 2543, 9626, 13231, 32409, 84260, 2882, 8203, 32678 OR
FILE REFERENCE: MPI03-007P1RNM0NIM
CURRENT APPLICATION NUMBER: US/10757, 262
CURRENT FILING DATE: 2004-01-14
PRIOR APPLICATION NUMBER: US 60/440, 318
PRIOR FILING DATE: 2003-01-15
PRIOR APPLICATION NUMBER: US 60/444, 783
PRIOR FILING DATE: 2003-02-04
PRIOR APPLICATION NUMBER: US 60/457, 901
PRIOR FILING DATE: 2003-03-27
PRIOR APPLICATION NUMBER: US 60/468, 775
PRIOR FILING DATE: 2003-05-08
PRIOR APPLICATION NUMBER: US 60/471, 614
PRIOR FILING DATE: 2003-05-19
PRIOR APPLICATION NUMBER: US 60/478, 742
PRIOR FILING DATE: 2003-06-16
PRIOR APPLICATION NUMBER: US 60/488, 529
PRIOR FILING DATE: 2003-07-18
PRIOR APPLICATION NUMBER: US 60/491, 156
PRIOR FILING DATE: 2003-07-30
PRIOR APPLICATION NUMBER: US 60/499, 594
PRIOR FILING DATE: 2003-09-02
PRIOR APPLICATION NUMBER: US 60/506, 332
PRIOR FILING DATE: 2003-09-26
NUMBER OF SEQ ID NOS: 136
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 52
LENGTH: 2053
TYPE: PRT
ORGANISM: Homo sapiens
US-10-757-262-52

Query Match 95.5%; Score 10022.5; DB 16; Length 2053;
Best Local Similarity 95.7%; Pred. No. 0;
Matches 1979; Conservative 3; Mismatches 5; Indels 81; Gaps 3;
Qy 1 MLKFKYGARBNPLDGAAPLARSRLNLFQKRPPTQOQMSPLSRBGILDLALFVLF 60


```

Db      1 MLKFKYGRBNPLDGAAPETASRSRLNLFPGKPPFTQOQMBPLSEGLIDLFLVLE 60
Qy      61 ECGPALMKIKHVSFVRKYSDDTIAELOPQSAKDEFVRSIVGCGHFAVOVVRKATG 120
Db      61 ECGPALMKIKHVSFVRKYSDDTIAELOPQSAKDEFVRSIVGCGHFAVOVVRKATG 120
Qy      121 DIYAMKMKKALLAOQVSPFEERINLSRSTSPWLOLAYAFODKXHLVLYMEYOPGG 180
Db      121 DIYAMKMKKALLAOQVSPFEERINLSRSTSPWLOLAYAFODKXHLVLYMEYOPGG 180
Qy      181 DLASLARVEDQDENLIFOYLAELILAVHSHVLMGVYHRDIKPBNTLVDTGHIKLVDF 240
Db      181 DLASLARVEDQDENLIFOYLAELILAVHSHVLMGVYHRDIKPBNTLVDTGHIKLVDF 240
Qy      241 GSAAKNSNKNVNAKPIGTDPYMAPEVLTVMNGDGKTYGLDGDWMSVGVIAVEMTYGR 300
Db      241 GSAAKNSNKNVNAKPIGTDPYMAPEVLTVMNGDGKTYGLDGDWMSVGVIAVEMTYGR 300
Qy      301 SPFAEGTSARTENNINMFORFLKFPDPKYSDDLIOQLCGQKRLKEGGLCHPFF 360
Db      301 SPFAEGTSARTENNINMFORFLKFPDPKYSDDLIOQLCGQKRLKEGGLCHPFF 360
Qy      361 SKIDMNNIRNSPPFVPTLKSDDDTSNPDEBKNSWSSSPCOLSPSGSGEELPFVGS 420
Db      361 SKIDMNNIRNSPPFVPTLKSDDDTSNPDEBKNSWSSSPCOLSPSGSGEELPFVGS 420
Qy      421 YSKALGILGRSESVVSGLDSPAKTSMEKLLIKSKELODQDCHKMEQMTLHRVS 480
Db      421 YSKALGILGRSESVVSGLDSPAKTSMEKLLIKSKELODQDCHKMEQMTLHRVS 480
Qy      481 EYEAVALSOKEVELKASTQSLBEOFLATYITTECSSIKRSLEQARMVESQEDKALQLH 540
Db      481 EYEAVALSOKEVELKASTQSLBEOFLATYITTECSSIKRSLEQARMVESQEDKALQLH 540
Qy      541 DIREOSRKLQEIKEOEYOAEVEMRLMNOLEEDLVABRRSDLYESELRESLAAEFK 600
Db      541 DIREOSRKLQEIKEOEYOAEVEMRLMNOLEEDLVABRRSDLYESELRESLAAEFK 600
Qy      601 RKAETECOHKLKAKDOQKPEVGEYAKLEKINAEQOLKIOELOELFAVASTATEBLQ 660
Db      601 RKAETECOHKLKAKDOQKPEVGEYAKLEKINAEQOLKIOELOELFAVASTATEBLQ 660
Qy      661 NIKOAKERARELEKLONRDSSSGIRKKLVBAERHSHLENKVKLETMERRENRLKD 720
Db      661 NIKOAKERARELEKLONRDSSSGIRKKLVBAERHSHLENKVKLETMERRENRLKD 720
Qy      721 IOTKSQOIQOMADKILEBEKHREAOVSQOHLBVHLKQKQOHTBEKIKVLDNOIKKDLAD 780
Db      721 IOTKSQOIQOMADKILEBEKHREAOVSQOHLBVHLKQKQOHTBEKIKVLDNOIKKDLAD 780
Qy      781 KETLENNMOQHEBEAHEKILISEOKAMINAMDSKIRLEBORIYELSEANKLAANSFLT 840
Db      781 KETLENNMOQHEBEAHEKILISEOKAMINAMDSKIRLEBORIYELSEANKLAANSFLT 840
Qy      841 ORNMKAQOEMIISLRQOKFYLETQAKLEAONRKLBOLEKISHODSDXNRLLLETRL 900
Db      841 ORNMKAQOEMIISLRQOKFYLETQAKLEAONRKLBOLEKISHODSDXNRLLLETRL 900
Qy      901 RETSLHEBEOKLEKQLTLOSLQJRESQLTNLQAAALASQLRQATTELEETTAA 960
Db      901 RETSLHEBEOKLEKQLTLOSLQJRESQLTNLQAAALASQLRQATTELEETTAA 960
Qy      961 EEBIOLATARDEIORKFDALRNSCTYITLBEOLNLTEDNALNNQNFYLSKOLDEAS 1020
Db      961 EEBIOLATARDEIORKFDALRNSCTYITLBEOLNLTEDNALNNQNFYLSKOLDEAS 1020
Qy      1005 GANDEIYOLASVDHLRREITEREMOJTSQKOTMEALKTTCMLBEQVMDLEALNDELLE 1080
Db      1005 GANDEIYOLASVDHLRREITEREMOJTSQKOTMEALKTTCMLBEQVMDLEALNDELLE 1080
Qy      1081 KEROWEAMRSVLDDEKSOFCRCRVELQOMLDTKQSAARADOKITTESROYVELAVXENKA 1140

```

```

Db      1065 KEROWEAMRSVLDDEKSOFCRCRVELQOMLDTKQSAARADOKITTESROYVELAVXENKA 1124
Qy      1141 EILALQALKEQKXKASLSDKLNDLEBKXAMLENNASLQOKLETTERELKORLLEQAK 1200
Db      1125 EILALQALKEQKXKASLSDKLNDLEBKXAMLENNASLQOKLETTERELKORLLEQAK 1184
Qy      1201 LOQOMDIQKXNIFULTOGLQBALRADLLKTERSDLEVOLENIQVLSHEKVRKEGTSIQ 1260
Db      1185 LOQOMDIQKXNIFULTOGLQBALRADLLKTERSDLEVOLENIQVLSHEKVRKEGTSIQ 1244
Qy      1261 QTKLIDFLQAKMDPAKXKX-----VPLQVNLKTLAEKAKCALEBEA 1305
Db      1245 QTKLIDFLQAKMDPAKXKXGKLSRRKEDPALPTQVLOVNLKTLAEKAKCALEBEA 1304
Qy      1306 LQKTRIELRSAREBAARAKATDHPHSTPATARQOIMSAIVRSPEHQPSAMSLAPSS 1365
Db      1305 LQKTRIELRSAREBAARAKATDHPHSTPATARQOIMSAIVRSPEHQPSAMSLAPSS 1364
Qy      1366 RRKESSTPEEFRRRLKERMHNIPHRNVGLNMPATCAVCLDVHFGROASKLECOVM 1425
Db      1365 RRKESSTPEEFRRRLKERMHNIPHRNVGLNMPATCAVCLDVHFGROASKLECOVM 1424
Qy      1426 CHPKCSTCLPATGCLPAEYATHFTEARCDPKMSPGLOTYEPSSSLHEGMMKVPNNKR 1485
Db      1425 CHPKCSTCLPATGCLPAEYATHFTEARCDPKMSPGLOTYEPSSSLHEGMMKVPNNKR 1484
Qy      1486 GQOGMDRKYIYLEGSKYLIYDNEAREAGORPVEEFELCPDGDVSIHGAVASSELAUTAK 1545
Db      1485 GQOGMDRKYIYLEGSKYLIYDNEAREAGORPVEEFELCPDGDVSIHGAVASSELAUTAK 1544
Qy      1546 ADVPYILKMSHPHTTCWPGRTIYLLAPSPDKQWVTALESVAVAGRVSRKADAKL 1605
Db      1545 A-----EKAEDAKL 1554
Qy      1606 LGNSILKLBDGDDLDNMCTLPFSQVVLVGTBEGYALNVLKNSLTHVPGGAVFOIYI 1665
Db      1555 LGNSILKLBDGDDLDNMCTLPFSQVVLVGTBEGYALNVLKNSLTHVPGGAVFOIYI 1614
Qy      1666 KDLEKLMINGEBRALCLVGVKKVQKSLAOSHLPACPDISPNTPEAVKGCILPAGAKTEN 1725
Db      1615 KDLEKLMINGEBRALCLVGVKKVQKSLAOSHLPACPDISPNTPEAVKGCILPAGAKTEN 1674
Qy      1726 GLCICAMPKSKVILARNENLSKXCIRKEIETSEBPCSHFTNYSILIGNKFEYIDMKQ 1785
Db      1675 GLCICAMPKSKVILARNENLSKXCIRKEIETSEBPCSHFTNYSILIGNKFEYIDMKQ 1734
Qy      1786 YTLBEFLDKNDHSLAPVFAASSNSFPVSIYOVNSAQOREBYLLCFHEFGVFTDYSGRS 1845
Db      1735 YTLBEFLDKNDHSLAPVFAASSNSFPVSIYOVNSAQOREBYLLCFHEFGVFTDYSGRS 1794
Qy      1846 RTDCLKMSRLPLAFAYBEPLYTFTHFNSLEVIETIOARSSAGTPARAVLDIPNRYLGPAT 1905
Db      1795 RTDCLKMSRLPLAFAYBEPLYTFTHFNSLEVIETIOARSSAGTPARAVLDIPNRYLGPAT 1864
Qy      1906 SSGAITYLASSYODKLRVICCKGNLVKESGTEHHRGPSTSRSSPNKRGPTYNHEITRVA 1965
Db      1855 SSGAITYLASSYODKLRVICCKGNLVKESGTEHHRGPSTSRSSPNKRGPTYNHEITRVA 1914
Qy      1966 SSPAPPEGPHREPESTPHRYRGRTYELARDKSPGRPLEKESKGRILSTRRESPARLF 2025
Db      1915 SSPAPPEGPHREPESTPHRYRGRTYELARDKSPGRPLEKESKGRILSTRRESPARLF 1974
Qy      2026 EDSSRGRPLPAGAVRTPLSOVNKAWDOS 2053
Db      1975 EDSSRGRPLPAGAVRTPLSOVNKAWDOS 2002

```

```

RESULT 12
US-10-028-946-4
; Sequence 4, Application US/10028946
; Publication NO. US20020123622A1
; GENERAL INFORMATION:
; APPLICANT: Yu, Xuanchuan

```

APPLICANT: Miranda, Maricar
APPLICANT: Fridtjof, Carl Johan
TITLE OF INVENTION: No. US20020123622A1 Human Kinases and Polynucleotides Encoding
FILE REFERENCE: LEX-0289-USA
CURRENT APPLICATION NUMBER: us/10/028,946
PRIOR FILING DATE: 2001-12-20
PRIOR APPLICATION NUMBER: US 60/258,335
NUMBER OF SEQ ID NOS: 4
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 4
LENGTH: 1958
TYPE: PRT
ORGANISM: homo sapiens
US-10-028-946-4

Query Match 94.5%; Score 9916; DB 13; Length 1958;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1945; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLKFKYGARNDLGAAPLPIASRSLNLFPGCKPMTQOQMSPLSREGILDALFVLFE 60
DB 1 MLKFKYGARNDLGAAPLPIASRSLNLFPGCKPMTQOQMSPLSREGILDALFVLFE 60
QY 61 ECGQPALMKIKHVSNFVRKYSDTTIAETQELQPSAKDFEVRSLVCGHFAEVQVREKATG 120
DB 61 ECGQPALMKIKHVSNFVRKYSDTTIAETQELQPSAKDFEVRSLVCGHFAEVQVREKATG 120
QY 121 DIYAMVKKKALLAOBOVSFPEEERNITLSRSTSPMIPOLQYAFODKNHLYLWMEYOPGG 180
DB 121 DIYAMVKKKALLAOBOVSFPEEERNITLSRSTSPMIPOLQYAFODKNHLYLWMEYOPGG 180
QY 181 DLSTLNRYEDQDENLQFYLAELILAVSHVLMGVHRDIKRENILVDRTGHIKLVD 240
DB 181 DLSTLNRYEDQDENLQFYLAELILAVSHVLMGVHRDIKRENILVDRTGHIKLVD 240
QY 241 GSAKAKNSNMKNVAKLPIGTPTDYAPFVLTVMNGDGKTYGLCDMWSVGIAYEMTYGR 300
DB 241 GSAKAKNSNMKNVAKLPIGTPTDYAPFVLTVMNGDGKTYGLCDMWSVGIAYEMTYGR 300
QY 301 SPBAEGSARTFNINMFORFLKPPDPKYSDFLDLQSLGCKEKLFEGLCGHPF 360
DB 301 SPBAEGSARTFNINMFORFLKPPDPKYSDFLDLQSLGCKEKLFEGLCGHPF 360
QY 361 SKIDMNNIRNSPPFPYTLKSDDDTSNFDPEKNSWVSSSPCOLSPSGFGEELPFVFGS 420
DB 361 SKIDMNNIRNSPPFPYTLKSDDDTSNFDPEKNSWVSSSPCOLSPSGFGEELPFVFGS 420
QY 421 YSKALGILGRSESVSGLDSPAKTSSMEKLLIKSKELQDSODKCHMEQEMTRLHRVS 480
DB 421 YSKALGILGRSESVSGLDSPAKTSSMEKLLIKSKELQDSODKCHMEQEMTRLHRVS 480
QY 481 EVAVALSOKVEVLKASTQSLLEODLATYITECSSIKRSLQEQARMVSDQEDKALQLH 540
DB 481 EVAVALSOKVEVLKASTQSLLEODLATYITECSSIKRSLQEQARMVSDQEDKALQLH 540
QY 541 DIRESQKLOEIKEOEQAOVBEKRLMMNQLLEBDLVASARRSDLYESELSESLAAEFK 600
DB 541 DIRESQKLOEIKEOEQAOVBEKRLMMNQLLEBDLVASARRSDLYESELSESLAAEFK 600
QY 601 RKATECOHKLKAKDQKPEVGEYAKLEKINAEOQLKIQELQELKAVASTEATELLQ 660
DB 601 RKATECOHKLKAKDQKPEVGEYAKLEKINAEOQLKIQELQELKAVASTEATELLQ 660
QY 661 NITQAKERARRELKQNRDSSSGITRKIVAEAEERHSLLENKYRLLETMERENRLKOD 720
DB 661 NITQAKERARRELKQNRDSSSGITRKIVAEAEERHSLLENKYRLLETMERENRLKOD 720
QY 721 IOTRSQOIQOAMDKILELEKREBAQVSAOHLLEVHLKQKQOYHEBKIKVDNQKDLAD 780
DB 721 IOTRSQOIQOAMDKILELEKREBAQVSAOHLLEVHLKQKQOYHEBKIKVDNQKDLAD 780
QY 781 KETLENNMQHBEAEHEKGLISEQKAMINAMDSKIRSLRQIVEISEANKLANSSFLT 840

DB 781 KETLENNMQHBEAEHEKGLISEQKAMINAMDSKIRSLRQIVEISEANKLANSSFLT 840
QY 841 ORNMKAQEMISLROQKLYLETOACKLEAONRKLEBEOLKISHODSDKNRLLEETRL 900
DB 841 ORNMKAQEMISLROQKLYLETOACKLEAONRKLEBEOLKISHODSDKNRLLEETRL 900
QY 901 REVSLHEBOKLEKQLTELQSLQERSQULTALQAAALLESQURQAKTELEETTAEA 960
DB 901 REVSLHEBOKLEKQLTELQSLQERSQULTALQAAALLESQURQAKTELEETTAEA 960
QY 961 EEBIOLATHRDIOKRPALRNSCTVITDIEBQLNQLTEBNALNNQNFYLSKQDEAS 1020
DB 961 EEBIOLATHRDIOKRPALRNSCTVITDIEBQLNQLTEBNALNNQNFYLSKQDEAS 1020
QY 1021 GANDEIVOLRSEVDHNRREITEREMQLTQKQTEALKTTCMLBEQVMDLEALNDELLE 1080
DB 1021 GANDEIVOLRSEVDHNRREITEREMQLTQKQTEALKTTCMLBEQVMDLEALNDELLE 1080
QY 1081 KERQWAMRSVLGDEKSOFECEVRBELQRMIDTEKQSRADQRTTESQVVELAVENHKA 1140
DB 1081 KERQWAMRSVLGDEKSOFECEVRBELQRMIDTEKQSRADQRTTESQVVELAVENHKA 1140
QY 1141 EITALLQALKEQKLKAESLSDKLNDELKHAMLEMMARSLOQLETERELKORLLEQAK 1200
DB 1141 EITALLQALKEQKLKAESLSDKLNDELKHAMLEMMARSLOQLETERELKORLLEQAK 1200
QY 1201 LQOQMDLOKNHIFRLTQGLQELADRADLKTERTSDLEYOLENTQVLSHEKVMEGTISQ 1260
DB 1201 LQOQMDLOKNHIFRLTQGLQELADRADLKTERTSDLEYOLENTQVLSHEKVMEGTISQ 1260
QY 1261 QTKLIDFLQAKMDQPAKKKVPLOYNELKLALKEKARCALEBALOKTRIELRSAREEA 1320
DB 1261 QTKLIDFLQAKMDQPAKKKVPLOYNELKLALKEKARCALEBALOKTRIELRSAREEA 1320
QY 1321 AHRKATDHPHSPAPARQOIMASATVRSPEHOPASMSLLAPSSSRKESSTPEESRRL 1380
DB 1321 AHRKATDHPHSPAPARQOIMASATVRSPEHOPASMSLLAPSSSRKESSTPEESRRL 1380
QY 1381 KERMHNNIPIHFPVNGLMRATKCAVCLDVTYHFGROASKLEQVMCHPKCSTGLPATCGL 1440
DB 1381 KERMHNNIPIHFPVNGLMRATKCAVCLDVTYHFGROASKLEQVMCHPKCSTGLPATCGL 1440
QY 1441 PAVYATHTBAFCRDYMSBGLQTKEPSSSLHEGMKTPRNKRGQOQDRKTYLLEGS 1500
DB 1441 PAVYATHTBAFCRDYMSBGLQTKEPSSSLHEGMKTPRNKRGQOQDRKTYLLEGS 1500
QY 1501 KVLIVYNEAREAGORVBEPELCLPDGVSIHGAVASLANTAKADVPIYILMESHPHT 1560
DB 1501 KVLIVYNEAREAGORVBEPELCLPDGVSIHGAVASLANTAKADVPIYILMESHPHT 1560
QY 1561 TCWPGRTLYLLAPSPDKORWYALLESVVAAGVRSREKADAKLGNLSLKLEGGDRULD 1620
DB 1561 TCWPGRTLYLLAPSPDKORWYALLESVVAAGVRSREKADAKLGNLSLKLEGGDRULD 1620
QY 1621 MNCTLPESDQVVLVGTBGLYALNVLKNSLTHVPGI GAVPQIYI IYDLEKLMIAEBERA 1680
DB 1621 MNCTLPESDQVVLVGTBGLYALNVLKNSLTHVPGI GAVPQIYI IYDLEKLMIAEBERA 1680
QY 1681 LCLVDYKVKYQSLAOSHLLAOPDISPNIPEAVKGCILFGAGKLENGLCI CAAPSPKVILL 1740
DB 1681 LCLVDYKVKYQSLAOSHLLAOPDISPNIPEAVKGCILFGAGKLENGLCI CAAPSPKVILL 1740
QY 1741 RYVENLSKTCIRKEIETSEPCSIHFTNYSILGTNKFETIDMKQYTLLEFDLKNDSILA 1800
DB 1741 RYVENLSKTCIRKEIETSEPCSIHFTNYSILGTNKFETIDMKQYTLLEFDLKNDSILA 1800
QY 1801 PAVFAASNSFPVSIQVNASAGREBYLLCFHEFGVFDVSYGRRSTDDLKMSRLPLAFA 1860
DB 1801 PAVFAASNSFPVSIQVNASAGREBYLLCFHEFGVFDVSYGRRSTDDLKMSRLPLAFA 1860
QY 1861 YREPPIYVTHFNSLEAYIELQARSACTPARAYLIDINPPLYGAISSGAIYLAASSVODKL 1920

Db 1861 YREPLYFVTHFNLSLEVEIEIQARSSAGTPARAYLIDIPNRYLGPALSSGATLYASSYODKL 1920
QY 1921 RVICCKGNLVESGTEHHRGPSTSR 1945
Db 1921 RVICCKGNLVESGTEHHRGPSTSR 1945

RESULT 13
US-09-964-956-40
Sequence 40, Application US/09964956
Publication No. US20040043926A1
GENERAL INFORMATION:
APPLICANT: Gerlach, Valerie L
APPLICANT: MacDougall, John R
APPLICANT: Smithson, Glenda
APPLICANT: Miller, Isabelle
APPLICANT: Stone, David
APPLICANT: Gunther, Erik
APPLICANT: Ellerman, Karen
APPLICANT: Grose, William M
APPLICANT: Alsobrook II, John P
APPLICANT: Lepley, Denise M
APPLICANT: Burgess, Catherine E
APPLICANT: Padigaru, Muralidhara
APPLICANT: Kekuda, Ramesh
APPLICANT: Spytek, Kimberly A
APPLICANT: Leach, Martin D
APPLICANT: Shukets, Richard A
TITLE OF INVENTION: No. US20040043926A1 Proteins and Nucleic Acids Encoding Same
FILE REFERENCE: 21402-124
CURRENT APPLICATION NUMBER: US/09/964,956
CURRENT FILING DATE: 2001-09-26
PRIOR APPLICATION NUMBER: 60/235,631
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: 60/235,633
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: 60/235,808
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: 60/236,064
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: 60/236,065
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: 60/236,066
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: 60/236,135
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: 60/236,135
PRIOR FILING DATE: 2000-09-28
PRIOR APPLICATION NUMBER: 60/237,434
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/238,321
PRIOR FILING DATE: 2000-10-05
PRIOR APPLICATION NUMBER: 60/238,399
PRIOR FILING DATE: 2000-10-06
PRIOR APPLICATION NUMBER: 60/238,396
PRIOR FILING DATE: 2000-10-06
PRIOR APPLICATION NUMBER: 60/276,667
PRIOR FILING DATE: 2001-03-16
PRIOR APPLICATION NUMBER: 60/294,823
PRIOR FILING DATE: 2001-05-31
PRIOR APPLICATION NUMBER: 60/304,868
PRIOR FILING DATE: 2001-07-12
NUMBER OF SEQ ID NOS: 127
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 40
LENGTH: 1641
TYPE: PRT
ORGANISM: Mus musculus
US-09-964-956-40

Query Match 77.8%; Score 8161; DB 11; Length 1641;
Best Local Similarity 95.7%; Pred. No. 0;
Matches 1610; Conservative 15; Mismatches 14; Indels 44; Gaps 2;

QY 374 PVPVTLKSDDDTSNFDPEKNSWVSSPCQLSPSGFGSEELPFVGFYSKALGILGRSES 433

Db 1 PVPVTLKSDDDTSNFDPEKNSWVSSPCQLSPSGFGSEELPFVGFYSKALGILGRSES 60
QY 434 VVSGLDSPAKTSSMEKLLIKSKELQDSODKCHMEQEMTTHRVSEVAVLQKEVEL 493
Db 61 VVSLDSPAIVSSMEKLLIKSKELQDSQDKCHMEQEMTTHRVSEVAVLQKEVEL 120
QY 494 KASETORSLLEODLATYITTECSSLKRSLEQARMEVSOEDDKALQLLIDREQSRLOEIK 553
Db 121 KASETORSLLEODLATYITTECSSLKRSLEQARMEVSOEDDKALQLLIDREQSRLOEIK 180
QY 554 EOEYQAOVEEMRLMMNOLEEDLVASARRSDLYSESLAESRLAAEFPRKATECQHKLLKA 613
Db 181 EOEYQAOVEEMRLMMNOLEEDLVASARRSDLYSESLAESRLAAEFPRKATECQHKLLKA 240
QY 614 KDOGPEVGEYAKLEKINAEQOLKIOLEKLEAVASTATELLOIROAKRAEREL 673
Db 241 KDOGPEVGEYAKLEKINAEQOLKIOLEKLEAVASTATELLOIROAKRAEREL 300
QY 674 EKLQNRDESGIRKCLVEAEERHSLBNKYKRLTEYERRENRKLDIQTSSQOIQOMAD 733
Db 301 EKLHNRDESGIRKCLVEAE----- 321
QY 734 KILEEKHREAOVSAGHLEVHLKQEQHYBEKIVLDNOIKKDLADKETLENNMOHRE 793
Db 322 ---ELEKHEHREAOVSAGHLEVHLKQEQHYBEKIVLDNOIKKDLADKETLENNMOHRE 378
QY 794 EAEHKGKILSEOKAMINAMDSKIRSLRQRYVELSEANKLANSSLFTORNNKAQEBMISE 853
Db 379 EAEHKGKILSEOKAMINAMDSKIRSLRQRYVELSEANKLANSSLFTORNNKAQEBMISE 438
QY 854 LRQCKFYLETOAGKLEAQNKRLEBOLEKISHODSDNRRLLETRLREYSLHEBOELE 913
Db 439 LRQCKFYLETOAGKLEAQNKRLEBOLEKISHODSDNRRLLETRLREYSLHEBOELE 498
QY 914 LKROLTELQSLQREESQTLALQAPRALSSQLQOATTELETAEABEELQALTARDE 973
Db 499 LKROLTELQSLQREESQTLALQAPRALSSQLQOATTELETAEABEELQALTARDE 558
QY 974 IORFQDALRNSCTVITLSEOLNLTEDNAELNNQNYLSKQLEASGANEIYQLRSEV 1033
Db 559 IORFQDALRNSCTVITLSEOLNLTEDNAELNNQNYLSKQLEASGANEIYQLRSEV 618
QY 1034 DHLREITEREMQUTSQOTMEALKTTCTMLEBOVMDLEALNDELKERQMEAMRSVLG 1093
Db 619 DHLREITEREMQUTSQOTMEALKTTCTMLEBOVMDLEALNDELKERQMEAMRSVLG 678
QY 1094 DEKSOFEQVRELOQMLDTEKOSPARADORTTESQYVELAVENKAEIILAQALKEQK 1153
Db 679 DEKSOFEQVRELOQMLDTEKOSPARADORTTESQYVELAVENKAEIILAQALKEQK 738
QY 1154 LKAEESLSDKNDLEKRAMLEMMARSLQKLETRERLQORLEQALQOQMDLQKNHIF 1213
Db 739 LKAEESLSDKNDLEKRAMLEMMARSLQKLETRERLQORLEQALQOQMDLQKNHIF 798
QY 1214 RLTOGLQEALDRADLKTERTSDLEYOLENIQVLYSHEKVMKEGTSQOTKLIDLEQAKMD 1273
Db 799 RLTOGLQEALDRADLKTERTSDLEYOLENIQVLYSHEKVMKEGTSQOTKLIDLEQAKMD 858
QY 1274 QPAKKKKVPLQYNELKLALKEKARCALEBEALQKTRIELRSAREBAHRKATDHPHST 1333
Db 859 QPAKKKKVPLQYNELKLALKEKARCALEBEALQKTRIELRSAREBAHRKATDHPHST 918
QY 1334 PATRQOIASATVRSSEHQPSAMSLAPSSSRKESSTPEFSRRLKERHNNHIPPFRN 1393
Db 919 PATRQOIASATVRSSEHQPSAMSLAPSSSRKESSTPEFSRRLKERHNNHIPPFRN 978
QY 1394 VGLNMRATKCAVCLDYTHFGROASKLEQVMCHPKSCSTCLPATCGLPAYATHTFAFC 1453
Db 979 VGLNMRATKCAVCLDYTHFGROASKLEQVMCHPKSCSTCLPATCGLPAYATHTFAFC 1038
QY 1454 RDKNSFGLOTKEPSSSLHLEGMKVPNNKRGQGWDRXYIVLSEKVLTYDNEARDAQ 1513

Db 1039 RDKNNSPGLQSKETGSSLIHLEGMKVPNNKRGQGGHMDRKITYLESGSKVLIYNNEAREAG 1098
Qy 1514 QRPVEEFELCLPDGDSVHIGAAGASBELANTAKADVPIYLKXESHPTTCTMPGRTLYLLAP 1573
Db 1099 QRPVEEFELCLPDGDSVHIGAAGASBELANTAKADVPIYLKXESHPTTCTMPGRTLYLLAP 1158
Qy 1574 SFPDKQWVYVNLBESVAVAGVRSEKAKADAKLGNLSLKLBGDRLLDMNCTLPSSDQVVL 1633
Db 1159 SFPDKQWVYVNLBESVAVAGVRSEKAKADAKLGNLSLKLBGDRLLDMNCTLPSSDQVVL 1218
Qy 1634 VGTBEGYALNVLKNSLTHVYVIGAVFOIYIYIKDLKELMAGBERALCLVDYKVKQSL 1693
Db 1219 VGTBEGYALNVLKNSLTHVYVIGAVFOIYIYIKDLKELMAGBERALCLVDYKVKQSL 1278
Qy 1694 AQSHLPAPQDIPSPNIFPAVYKCHLFGAKILENGLCIAMPSSKVYIIRYENISKCYCIRK 1753
Db 1279 AQSHLPAPQDIPSPNIFPAVYKCHLFGAKILENGLCIAMPSSKVYIIRYENISKCYCIRK 1338
Qy 1754 EIEFSEPCSCIHFTNYSILITGNTKFEYELDMQYTLDEFLDKNDHSLAPVPAASNSPVP 1813
Db 1339 EIEFSEPCSCIHFTNYSILITGNTKFEYELDMQYTLDEFLDKNDHSLAPVPAASNSPVP 1398
Qy 1814 SIYOVNAGOREEYILCFHEGVPVDSYGRSRTDLMKMSRLPLAFAYREBYLFTVTFNS 1873
Db 1399 SIYOVNAGOREEYILCFHEGVPVDSYGRSRTDLMKMSRLPLAFAYREBYLFTVTFNS 1458
Qy 1874 LEVLEIGARSSAGTAPAYALDIPNRYLGAIPAISSGAIYLAASYODKLVICCKGNLVKES 1933
Db 1459 LEVLEIGARSSAGTAPAYALDIPNRYLGAIPAISSGAIYLAASYODKLVICCKGNLVKES 1518
Qy 1934 GTEHHRPSTSRSPNKGSPPTVNEHTTKRVASPAPEGSHRREBTPHRY--REGRT 1991
Db 1519 GTEHHRPSTSRSPNKGSPPTVNEHTTKRVASPAPEGSHRREBTPHRYNDREGRT 1578
Qy 1992 ELRRDKSPGRLPESEKSPGRILSTRERSPARLPEDSRGRLLPAGAVRTPLSQVYKVMQD 2051
Db 1579 ELRRDKSPGRLPESEKSPGRILSTRERSPARLPEDSRGRLLPAGAVRTPLSQVYKVMQD 1638
Qy 2052 SSV 2054
Db 1639 SSV 1641

RESULT 14
US-10-017-216-5
; Sequence 5, Application US/10017216
; Publication No. US20020160483A1
; GENERAL INFORMATION:
; APPLICANT: KABELLER-LIBERMANN, Rosana
; TITLE OF INVENTION: 13245, A No. US20020160483A1e1 Human Myotonic Dystrophy Type Prot
; FILE REFERENCE: 10147-5701
; CURRENT APPLICATION NUMBER: US/10/017, 216
; CURRENT FILING DATE: 2001-10-23
; PRIOR APPLICATION NUMBER: US 60/242,429
; PRIOR FILING DATE: 2000-10-23
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 1641
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-017-216-5

Query Match 77.8%; Score 8161; DB 13; Length 1641;
Best Local Similarity 95.7%; Pred. No. 0;
Matches 1610; Conservative 15; Mismatches 14; Indels 44; Gaps 2;

Qy 374 PFVPTLKSDDDTSNFDPEPKNSWVSSPCQLSPGFGEBELPFVGFYSKALGILGRSES 433
Db 1 PFVPTLKSDDDTSNFDPEPKNSWVSSPCQLSPGFGEBELPFVGFYSKALGILGRSES 60
Qy 434 VVSGLDSPAKTSSMEKKLLIKSKELQDSQCKHMEQEMTRLLHRRVSEVAVALSQKEVEL 493

Db 61 VVSSLSPAKVSSMEKKLLIKSKELQDSQCKHMEQEMTRLLHRRVSEVAVALSQKEVEL 120
Qy 494 KASETORSLLEODLATYITECSSLKSLSEQARNEVSQEDDKALQLHDIREQSKLQEI 553
Db 121 KASETORSLLEODLATYITECSSLKSLSEQARNEVSQEDDKALQLHDIREQSKLQEI 180
Qy 554 EOEYQAOVEEMRLMNOLEBDLYSARRSDLYSELSRESLAEPEFRKATTEOQKILKA 613
Db 181 EOEYQAOVEEMRLMNOLEBDLYSARRSDLYSELSRESLAEPEFRKATTEOQKILKA 240
Qy 614 KDGQKPEVGEYAKLEKINAEOQLKIOELQKLEKAVASTATEATELLONIRQAKERAEREL 673
Db 241 KDGQKPEVGEYAKLEKINAEOQLKIOELQKLEKAVASTATEATELLONIRQAKERAEREL 300
Qy 674 EKLQNRDSSSEGIKKCLVEAEBRRHSLNKKRLFTMERRENRLKODIQKSOQIQOMAD 733
Db 301 EKLHNRDSSSEGIKKCLVEAEBRRHSLNKKRLFTMERRENRLKODIQKSOQIQOMAD 321
Qy 734 KIIELEBKREAOVSAQHLEVHLKQKQHYBEKIKYLDNOIKKDLADKETLENMQRHEE 793
Db 322 ---ELEBKREAOVSAQHLEVHLKQKQHYBEKIKYLDNOIKKDLADKETLENMQRHEE 378
Qy 794 EAEHKGKILSEOKAMINAMDSKIRSLFORIVELSEANKLAANSLLFTORMKAOEEMISE 853
Db 379 EAEHKGKILSEOKAMINAMDSKIRSLFORIVELSEANKLAANSLLFTORMKAOEEMISE 438
Qy 854 LRQOKFYLETOAKLEAONKRLBEOLKISHODSDKNRLLELETRLREVSLEHEOKLE 913
Db 439 LRQOKFYLETOAKLEAONKRLBEOLKISHODSDKNRLLELETRLREVSLEHEOKLE 498
Qy 914 LRQOKFYLETOAKLEAONKRLBEOLKISHODSDKNRLLELETRLREVSLEHEOKLE 973
Db 499 LRQOKFYLETOAKLEAONKRLBEOLKISHODSDKNRLLELETRLREVSLEHEOKLE 558
Qy 974 IORPPALNRSCVTITDLEBQNLQTEBDNAELNNOFYLISKQDEASGANDEIVOLSEV 1033
Db 559 IORPPALNRSCVTITDLEBQNLQTEBDNAELNNOFYLISKQDEASGANDEIVOLSEV 618
Qy 1034 DHLRREITEREMQJTSQKQTEALKTTCMLBEQVMDLEALNDELLEKEROEAMRSVLG 1093
Db 619 DHLRREITEREMQJTSQKQTEALKTTCMLBEQVMDLEALNDELLEKEROEAMRSVLG 678
Qy 1094 DEKSOPEBCRVREIQRMLDTEKOSARADQRTSSROYVELAVENHAKETIALLQALKEOK 1153
Db 679 DEKSOPEBCRVREIQRMLDTEKOSARADQRTSSROYVELAVENHAKETIALLQALKEOK 738
Qy 1154 LKAEISLSDKLNDELKKNAMLENNARSLOQLFETREIKORLLEBOAKLQOQMDLQKNHIF 1213
Db 739 LKAEISLSDKLNDELKKNAMLENNARSLOQLFETREIKORLLEBOAKLQOQMDLQKNHIF 798
Qy 1214 RLTOGLQEALDRADLLKTERSDLEYOLENIQVLYSHKVMMEGTISQQTGLIDFLQAKMD 1273
Db 799 RLTOGLQEALDRADLLKTERSDLEYOLENIQVLYSHKVMMEGTISQQTGLIDFLQAKMD 858
Qy 1274 QPAPKKKKVPLQVYELKALAEKAKACAELEALQKTRIELRSAREBAHKAADHHPST 1333
Db 859 QPAPKKKKVPLQVYELKALAEKAKACAELEALQKTRIELRSAREBAHKAADHHPST 918
Qy 1334 PATARQOIAMSAIVRSEPHOPSAISLLAPPSRRKSSSTPEERSRLKERMHNIPIHFN 1393
Db 919 PATARQOIAMSAIVRSEPHOPSAISLLAPPSRRKSSSTPEERSRLKERMHNIPIHFN 978
Qy 1394 VGLNMBATKCAVCLDVHFGROASKLCEQVMCHPKCSTCLPATCGLPAYATHPTFAFC 1453
Db 979 VGLNMBATKCAVCLDVHFGROASKLCEQVMCHPKCSTCLPATCGLPAYATHPTFAFC 1038
Qy 1454 RDKNNSPGLQSKETGSSLIHLEGMKVPNNKRGQGGHMDRKITYLESGSKVLIYNNEAREAG 1513
Db 1039 RDKNNSPGLQSKETGSSLIHLEGMKVPNNKRGQGGHMDRKITYLESGSKVLIYNNEAREAG 1098
Qy 1514 QRPVEEFELCLPDGDSVHIGAAGASBELANTAKADVPIYLKXESHPTTCTMPGRTLYLLAP 1573

Db 1099 QRPVEEELCLPDDVSIHGAVGASELANTAKADVPIILKNEHPHTTCWREPTLYLLAP 1158
 Qy 1574 SPFDKQWVTLASVAVAGRVSRKAEADATLGNLSLKJEGDRLDMNCTLPSPDQVYL 1633
 Db 1159 SFPDQKQWVTLASVAVAGRVSRKAEADATLGNLSLKJEGDRLDMNCTLPSPDQVYL 1218
 Qy 1634 VGTGEGYALNVLKNSLTHVPGICAVPQIYIILKLEKLMJAGEBRALCLVDYKVKVQSL 1633
 Db 1219 VGTGEGYALNVLKNSLTHVPGICAVPQIYIILKLEKLMJAGEBRALCLVDYKVKVQSL 1278
 Qy 1694 AQSHLPAPDISPNIFFAVKGCGLFGAGKINGLCICGAMPKXVILRYNENLSKYCIRK 1753
 Db 1279 AQSHLPAPDISPNIFFAVKGCGLFGAGKINGLCICGAMPKXVILRYNENLSKYCIRK 1338
 Qy 1754 EIEFSEPCSCIHFTNYSILIGTNKFEYIDMKQYTLBEFLDNDSHSLPAVFAASSNSFPV 1813
 Db 1339 EIEFSEPCSCIHFTNYSILIGTNKFEYIDMKQYTLBEFLDNDSHSLPAVFAASSNSFPV 1398
 Qy 1814 SIYQVNSAGQREBYLTCHEHGVFVDSYGRSRFTDLMKWSHLPLAFVREBYLFTVTHNS 1873
 Db 1399 SIYQVNSAGQREBYLTCHEHGVFVDSYGRSRFTDLMKWSHLPLAFVREBYLFTVTHNS 1458
 Qy 1874 LEVEIQARSSAGTPARAYLDIPNRYLGPALISSGATYLAASSYODKLRYLCCCKNLVYES 1933
 Db 1459 LEVEIQARSSAGTPARAYLDIPNRYLGPALISSGATYLAASSYODKLRYLCCCKNLVYES 1518
 Qy 1934 GTEHHRGPTSRSSPNKRGPTVNEHITTKRVASSPAPPEGSHPRESTPHRY--REGRT 1991
 Db 1519 GTEHHRGPTSRSSPNKRGPTVNEHITTKRVASSPAPPEGSHPRESTPHRY--REGRT 1578
 Qy 1992 ELRRDQSPGRPLEERKSPGRILSTRERSPARLFEDESSRGLPLGAVRTPLSQVKNYWDQ 2051
 Db 1579 ELRRDQSPGRPLEERKSPGRILSTRERSPARLFEDESSRGLPLGAVRTPLSQVKNYWDQ 1638
 Qy 2052 SSV 2054
 Db 1639 SSV 1641

RESULT 15

US-09-964-956-41
 ; Sequence 41, Application US/09964956
 ; Publication No. US20040043926A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Gerlach, Valerie L
 ; APPLICANT: MacDougall, John R
 ; APPLICANT: Smithson, Glenda
 ; APPLICANT: Miller, Isabelle
 ; APPLICANT: Stone, David
 ; APPLICANT: Gunther, Erik
 ; APPLICANT: Billeman, Karen
 ; APPLICANT: Grosbeck, William M
 ; APPLICANT: Alsobrook II, John P
 ; APPLICANT: Lepley, Denise M
 ; APPLICANT: Burgess, Catherine E
 ; APPLICANT: Padigaru, Muralidhara
 ; APPLICANT: Kekuda, Ramesh
 ; APPLICANT: Spytek, Kimberly A
 ; APPLICANT: Leach, Martin D
 ; APPLICANT: Shinkets, Richard A
 ; TITLE OF INVENTION: No. US20040043926A1 Proteins and Nucleic Acids Encoding Same
 ; FILE REFERENCE: 21402-124
 ; CURRENT APPLICATION NUMBER: US/09/964,956
 ; PRIOR FILING DATE: 2001-09-26
 ; PRIOR APPLICATION NUMBER: 60/235,631
 ; PRIOR FILING DATE: 2000-09-27
 ; PRIOR APPLICATION NUMBER: 60/235,633
 ; PRIOR FILING DATE: 2000-09-27
 ; PRIOR APPLICATION NUMBER: 60/235,808
 ; PRIOR FILING DATE: 2000-09-27
 ; PRIOR APPLICATION NUMBER: 60/236,064
 ; PRIOR FILING DATE: 2000-09-27
 ; PRIOR APPLICATION NUMBER: 60/236,065

; PRIOR FILING DATE: 2000-09-27
 ; PRIOR APPLICATION NUMBER: 60/236,066
 ; PRIOR FILING DATE: 2000-09-27
 ; PRIOR APPLICATION NUMBER: 60/236,135
 ; PRIOR FILING DATE: 2000-09-28
 ; PRIOR APPLICATION NUMBER: 60/237,434
 ; PRIOR FILING DATE: 2000-10-03
 ; PRIOR APPLICATION NUMBER: 60/238,321
 ; PRIOR FILING DATE: 2000-10-05
 ; PRIOR APPLICATION NUMBER: 60/238,399
 ; PRIOR FILING DATE: 2000-10-06
 ; PRIOR APPLICATION NUMBER: 60/238,396
 ; PRIOR FILING DATE: 2000-10-06
 ; PRIOR APPLICATION NUMBER: 60/276,667
 ; PRIOR FILING DATE: 2001-03-16
 ; PRIOR APPLICATION NUMBER: 60/294,823
 ; PRIOR FILING DATE: 2001-05-31
 ; PRIOR APPLICATION NUMBER: 60/304,868
 ; PRIOR FILING DATE: 2001-07-12
 ; NUMBER OF SEQ ID NOS: 127
 ; SOFTWARE: Patentin Ver. 2.1
 ; SEQ ID NO 41
 ; LENGTH: 1597
 ; TYPE: PRT
 ; ORGANISM: Mus musculus
 US-09-964-956-41
 Query Match 75.6%; Score 7935; DB 11; Length 1597;
 Best Local Similarity 98.2%; Pred. No. 0;
 Matches 1561; Conservative 16; Mismatches 10; Indels 2; Gaps 1;
 Qy 468 MEQEMTLHRRVSEVAVLSQKEVELKASFTORSLEBODLATYITEGSSLSRSLQARME 527
 Db 9 MEQEMTLHRRVSEVAVLSQKEVELKASFTORSLEBODLATYITEGSSLSRSLQARME 68
 Qy 528 VSQEDDYALQLHDIREQSRKLEIKQEYQAOVEERLMMNQLBEDLVARRRSDLYES 587
 Db 69 VSQEDDYALQLHDIREQSRKLEIKQEYQAOVEERLMMNQLBEDLVARRRSDLYES 128
 Qy 588 ELRESRLAAEFKRAKTECOHKLKAKDQKPEVGEYAKLEKINABOLKIOELQELK 647
 Db 129 ELRESRLAAEFKRAKTECOHKLKAKDQKPEVGEYAKLEKINABOLKIOELQELK 188
 Qy 648 AVKASTATELLQNIIRAKERAERELEKLNREDSSEGRKCLVEAERRSLENKYRL 707
 Db 189 AVKASTATELLQNIIRAKERAERELEKLNREDSSEGRKCLVEAERRSLENKYRL 248
 Qy 708 ETWERERRLKDDIQTSQOIQQWADKILEEKGREAOVSAOHLVHLKQEQHYEKI 767
 Db 249 ETWERERRLKDDIQTSQOIQQWADKILEEKGREAOVSAOHLVHLKQEQHYEKI 308
 Qy 768 KVLDDNICKDLAKETLEMMQHEERAAHKGKILSOKMINMDSKISLEQIVELS 827
 Db 309 KVLDDNICKDLADESLEMMQHEERAAHKGKILSOKMINMDSKISLEQIVELS 368
 Qy 828 EANKLAANSLFTQRRNKAAQEMISELRQOKFYLETAGKLEAONRKLEBOLKISHODH 887
 Db 369 EANKLAANSLFTQRRNKAAQEMISELRQOKFYLETAGKLEAONRKLEBOLKISHODH 428
 Qy 888 SDKRLILEETRLREVSLEHEBQKLEIKROLTEIQLSLQERESQTLQAARALLESQLR 947
 Db 429 SDKRLILEETRLREVSLEHEBQKLEIKROLTEIQLSLQERESQTLQAARALLESQLR 488
 Qy 948 QAKTELEETTAAREEETQALTARDEIQRPDALRNSCTVITDEBOLNQTENNALNN 1007
 Db 489 QAKTELEETTAAREEETQALTARDEIQRPDALRNSCTVITDEBOLNQTENNALNN 548
 Qy 1008 QNFYLSKQDASGANDIVQLRSEVDHLRREITEREMQLTQKQTEALKTCTMLBEQ 1067
 Db 549 QNFYLSKQDASGANDIVQLRSEVDHLRREITEREMQLTQKQTEALKTCTMLBEQ 608
 Qy 1068 VMDLEALNDELLEKERQEWARSVLGDEKSQFECRVRELQRMIDTEKQSPARADQRTTES 1127
 ; :

Db 609 VLDELINDELLEKEROMEAMRSVLGDBKSQFECRVRELOQMLDTEKQSRARADQRTTES 668

Qy 1128 ROVELLAVKEHKAETLALQALKEQKLKAESLSPKLNDELKHAMLENNARSLOQKLETE 1187

Db 669 ROVELLAVKEHKAETLALQALKEQKLKAESLSPKLNDELKHAMLENNARSLOQKLETE 728

Qy 1188 RELKORLLEBQAKLOQOMDLOKNHIFRLTOGLQALDRAJDLKTERSDLEYOLENIQVLY 1247

Db 729 RELKORLLEBQAKLOQOMDLOKNHIFRLTOGLQALDRAJDLKTERSDLEYOLENIQVLY 788

Qy 1248 SHEKVMKEGTISQOTKLIDFLQAMDOPAKKKKVPLQYNEIKLALKEKARCALIEBALQ 1307

Db 789 SHEKVMKEGTISQOTKLIDFLQAMDOPAKKKKVPLQYNEIKLALKEKARCALIEBALQ 848

Qy 1308 KTRIELRSAREEAHRRATDHPHPSTPATARQOIAMSAIVRSPEHOPSAMSLAPSSRR 1367

Db 849 KTRIELRSAREEAHRRATDHPHPSTPATARQOIAMSAIVRSPEHOPSAMSLAPSSRR 908

Qy 1368 KESTPEEFSGRLKERHNNIPHRFNVLNNRATKCAVCLDTVHFRQASKLECOVMCH 1427

Db 909 KESTPEEFSGRLKERHNNIPHRFNVLNNRATKCAVCLDTVHFRQASKLECOVMCH 968

Qy 1428 PKCSTCLPATCGLPAEYATHTFAFCRDKNNSPGLQTKESPSSSLHLEGMMKVPRNNKRGQ 1487

Db 969 PKCSTCLPATCGLPAEYATHTFAFCRDKNNSPGLQTKESPSSSLHLEGMMKVPRNNKRGQ 1028

Qy 1488 QGMWRKTYILEGSKVLYDNEAREAGORPVDEFELCLPDGDVSJHGA VGASELANTAKAD 1547

Db 1029 QGMWRKTYILEGSKVLYDNEAREAGORPVDEFELCLPDGDVSJHGA VGASELANTAKAD 1088

Qy 1548 VPYILKMSHPHTTQWGRITLYLLAPSPDKQWVTALLESVAGRVSRKAEADAKLIG 1607

Db 1089 VPYILKMSHPHTTQWGRITLYLLAPSPDKQWVTALLESVAGRVSRKAEADAKLIG 1148

Qy 1608 NSLLKLEGGDRDLDMNCTLPFSDQVVLVGTBEGLYALNVLKNSLTHVPGIGAFQIYIYIKD 1667

Db 1149 NSLLKLEGGDRDLDMNCTLPFSDQVVLVGTBEGLYALNVLKNSLTHVPGIGAFQIYIYIKD 1208

Qy 1668 LEKLIAMAGEERALCLVDVKKVKQSLAQSHLPAQPDISPNI FEAVKGCCHLFGAGXIENGL 1727

Db 1209 LEKLIAMAGEERALCLVDVKKVKQSLAQSHLPAQPDISPNI FEAVKGCCHLFGAGXIENGL 1268

Qy 1728 CIGAMPSKVILPYNMENTSKYCIKKEIETSEPCSIHFTMYSTLIGTNKFEYIDMKOYT 1787

Db 1269 CIGAMPSKVILPYNMENTSKYCIKKEIETSEPCSIHFTMYSTLIGTNKFEYIDMKOYT 1328

Qy 1788 LEEFLDKNDHSLAPVFAASNSFPVSIQVNSAGQREBYLLCFHEFGVFVDSYGRSRT 1847

Db 1329 LEEFLDKNDHSLAPVFAASNSFPVSIQVNSAGQREBYLLCFHEFGVFVDSYGRSRT 1388

Qy 1848 DDLKMSRLPLAFAYREPYLLPVTHNSLLEVIEIQARSSAGTPARAYLDIPNRYLGAIISS 1907

Db 1389 DDLKMSRLPLAFAYREPYLLPVTHNSLLEVIEIQARSSAGTPARAYLDIPNRYLGAIISS 1448

Qy 1908 GATYLAASYQDKLAVTCCKGLVKESTGTEHHRGSTRSSPNKRGPTNYNEHTKRYASS 1967

Db 1449 GATYLAASYQDKLAVTCCKGLVKESTGTEHHRGSTRSSPNKRGPTNYNEHTKRYASS 1508

Qy 1968 PAPPEGSHPREPSTPHRY--REGTELRDQSPGRPLERKSPGRITSTRERSPARLF 2025

Db 1509 PAPPEGSHPREPSTPHRYDRREKTELRDQSPGRPLERKSPGRITSTRERSPARLF 1568

Qy 2026 EDSGRGLPAGAVRTPLSQVNVKWDQSSV 2054

Db 1569 EDSGRGLPAGAVRTPLSQVNVKWDQSSV 1597

Search completed: May 11, 2005, 17:00:24
Job time : 215 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 11, 2005, 16:33:19 ; Search time 56 Seconds
(without alignments)
2738.021 Million cell updates/sec

Title: US-10-791-666-2
Perfect score: 10493
Sequence: 1 MLKFKYGARNPDLGAAAEPI.....AGAVRTPLSQVKNKWDQSSV 2054

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/ptodata/1/1aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/1/1aa/5B.COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/6A.COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/6B.COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/PCUTS.COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	10493	100.0	2054	US-10-028-946-2	Sequence 2, Appli
2	9916	94.5	1986	US-10-028-946-4	Sequence 4, Appli
3	2440	23.3	497	US-09-804-471A-2	Sequence 2, Appli
4	2440	23.3	497	US-10-238-709-2	Sequence 2, Appli
5	2175.5	20.7	494	US-09-804-471A-4	Sequence 4, Appli
6	2175.5	20.7	494	US-10-238-709-4	Sequence 4, Appli
7	1287	12.3	257	US-09-916-204-2	Sequence 2, Appli
8	1287	12.3	257	US-10-282-048-2	Sequence 2, Appli
9	1263.5	12.0	1386	US-08-685-576-1	Sequence 1, Appli
10	1260	12.0	1354	US-08-685-576-2	Sequence 1, Appli
11	1257	12.0	1386	US-08-685-576-4	Sequence 4, Appli
12	1257	12.0	1386	US-09-976-594-296	Sequence 26, App
13	1183	11.3	251	US-09-916-204-4	Sequence 4, Appli
14	1183	11.3	251	US-10-282-048-4	Sequence 4, Appli
15	1180	11.2	251	US-09-916-204-5	Sequence 5, Appli
16	1180	11.2	251	US-09-916-204-6	Sequence 6, Appli
17	1180	11.2	251	US-10-282-048-5	Sequence 5, Appli
18	1180	11.2	251	US-10-282-048-6	Sequence 6, Appli
19	1048.5	10.0	900	US-08-630-822A-62	Sequence 62, Appli
20	1048.5	10.0	900	US-09-005-069-62	Sequence 62, Appli
21	1048.5	10.0	900	US-09-171-156A-21	Sequence 21, Appli
22	1048.5	10.0	900	US-09-004-730A-21	Sequence 21, Appli
23	1048.5	10.0	900	US-08-981-799A-21	Sequence 21, Appli
24	906.5	8.6	509	US-09-949-016-8511	Sequence 8511, Ap
25	857.5	8.2	420	US-08-685-871-58	Sequence 58, Appli
26	853.5	8.1	582	US-08-422-699A-9	Sequence 9, Appli
27	853.5	8.1	582	US-08-422-706B-9	Sequence 9, Appli

28	840.5	8.0	420	3	US-08-685-871-59	Sequence 59, Appli
29	809.5	7.7	638	2	US-08-422-699A-11	Sequence 11, Appli
30	809.5	7.7	638	2	US-08-422-706B-11	Sequence 11, Appli
31	786	7.5	555	1	US-08-484-044-6	Sequence 6, Appli
32	715.5	6.8	479	3	US-09-442-100-13	Sequence 13, Appli
33	715.5	6.8	479	3	US-08-939-106-13	Sequence 13, Appli
34	715.5	6.8	479	4	US-09-442-102-13	Sequence 13, Appli
35	701	6.7	526	3	US-09-442-100-12	Sequence 12, Appli
36	701	6.7	526	4	US-08-939-106-12	Sequence 12, Appli
37	701	6.7	526	4	US-09-442-102-12	Sequence 12, Appli
38	686.5	6.5	719	3	US-09-588-256-2	Sequence 3, Appli
39	683	6.5	404	2	US-08-860-150-3	Sequence 3, Appli
40	683	6.5	404	3	US-09-338-132-3	Sequence 3, Appli
41	677	6.5	464	2	US-08-878-989-4	Sequence 4, Appli
42	677	6.5	464	3	US-09-272-796-4	Sequence 4, Appli
43	674.5	6.4	465	2	US-08-878-989-18	Sequence 18, Appli
44	674.5	6.4	465	2	US-08-860-150-7	Sequence 7, Appli
45	674.5	6.4	465	3	US-09-338-132-7	Sequence 7, Appli

ALIGNMENTS

RESULT 1
US-10-028-946-2
; Sequence 2, Application US/10028946
; Patent No. 6734009
; GENERAL INFORMATION:
; APPLICANT: Yu, Xuanchuan
; APPLICANT: Miranda, Maricar
; APPLICANT: Fiddale, Carl Ushan
; TITLE OF INVENTION: No. 6734009e1 Human Kinases and Polynucleotides Encoding the Same
; FILE REFERENCE: LEX-0289-USA
; CURRENT APPLICATION NUMBER: US/10/028,946
; CURRENT FILING DATE: 2001-12-20
; PRIOR APPLICATION NUMBER: US 60/258,335
; PRIOR FILING DATE: 2000-12-27
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 2054
; TYPE: PRT
; ORGANISM: homo sapiens
US-10-028-946-2

Query Match	Score	100.0%	DB 4	Length	2054
Best Local Similarity	100.0%	Pred. No. 0	Matches 2054	Conservative 0	Mismatches 0
Indels	0	Gaps	0		
QY	1	MLKFKYGARNPDLGAAAEPIASRSARLNTFFQKPPMTQOQNSPLSREGILDALFVLF	60		
DB	1	MLKFKYGARNPDLGAAAEPIASRSARLNTFFQKPPMTQOQNSPLSREGILDALFVLF	60		
QY	61	ECSPALMKIKHSNPFVKYSDTIAEQLOPSAKDFEVSILVCGHFAEYQVREKATG	120		
DB	61	ECSPALMKIKHSNPFVKYSDTIAEQLOPSAKDFEVSILVCGHFAEYQVREKATG	120		
QY	121	DIYAMKMKKKKALLAQOVSFFFEERNILSRSTSPWIPQIYARQDKNHLTVNVEYPGG	180		
DB	121	DIYAMKMKKKKALLAQOVSFFFEERNILSRSTSPWIPQIYARQDKNHLTVNVEYPGG	180		
QY	181	DLASLNRYEDQDENLQFYLAELIAYSVHLMGVYHDIKPEKENTLVRTGHIKLVDF	240		
DB	181	DLASLNRYEDQDENLQFYLAELIAYSVHLMGVYHDIKPEKENTLVRTGHIKLVDF	240		
QY	241	GSAAKNSNKNVNAKLPIDGPDYNAPEVLTVNMGDKGTGGLDCDMSVGVIAYEMTYGR	300		
DB	241	GSAAKNSNKNVNAKLPIDGPDYNAPEVLTVNMGDKGTGGLDCDMSVGVIAYEMTYGR	300		
QY	301	SPPAEGTSATFNINMFORFLKPPDPKXSSDFLDIOSLLCGQKRLKFBGICCHPFF	360		
DB	301	SPPAEGTSATFNINMFORFLKPPDPKXSSDFLDIOSLLCGQKRLKFBGICCHPFF	360		

QY 361 SKIDMNNIRNSPPFVETLKSDDTNSPNDPEKNSWSSSPSCQUSPFGSGEEELPVGFS 420
 DB 361 SKIDMNNIRNSPPFVETLKSDDTNSPNDPEKNSWSSSPSCQUSPFGSGEEELPVGFS 420
 QY 421 YSRALGILGSESVVGLSDPAKTSMEKLLIISKRLQSDQCHMEQEMRLHRRVS 480
 DB 421 YSRALGILGSESVVGLSDPAKTSMEKLLIISKRLQSDQCHMEQEMRLHRRVS 480
 QY 481 EBAVALISQKEVELKASETORSILEQDLATYITTECSSIKRSLEQARMEVSQEDDQLLH 540
 DB 481 EBAVALISQKEVELKASETORSILEQDLATYITTECSSIKRSLEQARMEVSQEDDQLLH 540
 QY 541 DIREQSRKQEIKEOEYOAQVEEMRLMNOLEEDLVASARRSDIYESELRESRLAAEFK 600
 DB 541 DIREQSRKQEIKEOEYOAQVEEMRLMNOLEEDLVASARRSDIYESELRESRLAAEFK 600
 QY 601 RKAITECHKLLKADQKPEVGEYAKLEKINAEQOLKIQELQELKAVASTATELLQ 660
 DB 601 RKAITECHKLLKADQKPEVGEYAKLEKINAEQOLKIQELQELKAVASTATELLQ 660
 QY 661 NIKQAKRAREBELEKLNREDSSEGRKKLVEABERHSLENKYRLETMERRENRLKDD 720
 DB 661 NIKQAKRAREBELEKLNREDSSEGRKKLVEABERHSLENKYRLETMERRENRLKDD 720
 QY 721 IQTKSQOIQWADKILELEEKREAOVSAQHLVHLKQEQHYEKKIYLDNQIKKDLAD 780
 DB 721 IQTKSQOIQWADKILELEEKREAOVSAQHLVHLKQEQHYEKKIYLDNQIKKDLAD 780
 QY 781 KETLENNMORHEEBAHKKILSFOKAMINAMDSKISLSERIVELSEANKLANSSIFT 840
 DB 781 KETLENNMORHEEBAHKKILSFOKAMINAMDSKISLSERIVELSEANKLANSSIFT 840
 QY 841 QRMKAQOEMISELROQFYLETQAGKLEAQNRLKEBOLKESISHQSDKRLLEETRL 900
 DB 841 QRMKAQOEMISELROQFYLETQAGKLEAQNRLKEBOLKESISHQSDKRLLEETRL 900
 QY 901 REVSLEHEBQKLEIKROLTELQLSLOERESQITLQARAPALBESQLOAKTELEETTAEA 960
 DB 901 REVSLEHEBQKLEIKROLTELQLSLOERESQITLQARAPALBESQLOAKTELEETTAEA 960
 QY 961 EEEIQAALTARDEIQRFEDALRNSCTVITDLEBOLNQTENALNQNPFILSQDLBAS 1020
 DB 961 EEEIQAALTARDEIQRFEDALRNSCTVITDLEBOLNQTENALNQNPFILSQDLBAS 1020
 QY 1021 GANDEIQVLRSEVVDLREITEREMOJTSOKQTEALKTCTMLBEQVMDLEALDELLE 1080
 DB 1021 GANDEIQVLRSEVVDLREITEREMOJTSOKQTEALKTCTMLBEQVMDLEALDELLE 1080
 QY 1081 KERQWEAMRSVLDGKESQFECRAVELQRMIDTERKQSPARADQRTTESRQVVELAVKEHKA 1140
 DB 1081 KERQWEAMRSVLDGKESQFECRAVELQRMIDTERKQSPARADQRTTESRQVVELAVKEHKA 1140
 QY 1141 EILALQALKEQKIKASLSLNDLEKKAHLEMMARSLQOKLETERELKORLLEEQAK 1200
 DB 1141 EILALQALKEQKIKASLSLNDLEKKAHLEMMARSLQOKLETERELKORLLEEQAK 1200
 QY 1201 LQOQMDLOKNHIFLTLQLOEALRADLKTTERSDFLETOLENIQVLSHEKVKMEGTISQ 1260
 DB 1201 LQOQMDLOKNHIFLTLQLOEALRADLKTTERSDFLETOLENIQVLSHEKVKMEGTISQ 1260
 QY 1261 QTKLIDFLQAKMDQPAKKKKVPLQYNELKALKEKARCABELEALQKTRITELASABEA 1320
 DB 1261 QTKLIDFLQAKMDQPAKKKKVPLQYNELKALKEKARCABELEALQKTRITELASABEA 1320
 QY 1321 AHRKATHPHPSTPATARQOJAMSAIVRSPEHQSSANSLAPSSRRKESSTPEEFRRRL 1380
 DB 1321 AHRKATHPHPSTPATARQOJAMSAIVRSPEHQSSANSLAPSSRRKESSTPEEFRRRL 1380
 QY 1381 KERHNNHIPIHPFNVLGMBATKCAVCLDTYVHFGQASCLCEQVMCHPKCSTCLPATCGL 1440
 DB 1381 KERHNNHIPIHPFNVLGMBATKCAVCLDTYVHFGQASCLCEQVMCHPKCSTCLPATCGL 1440
 QY 1441 PAEYATHTFAFCRDKNRNSPGLQTKPESSSLHLEGMKMKVPRNNRGOQGMGRKYIVLEGS 1500

DB 1441 PAEYATHTFAFCRDKNRNSPGLQTKPESSSLHLEGMKMKVPRNNRGOQGMGRKYIVLEGS 1500
 QY 1501 KYLIDYNEAREAGQRPVEBEFLCLPDGDVSIHGAVGASELANTAKADVPIYLXNESHPT 1560
 DB 1501 KYLIDYNEAREAGQRPVEBEFLCLPDGDVSIHGAVGASELANTAKADVPIYLXNESHPT 1560
 QY 1561 TCMPGRTLYLLASPDPKQKRWATLALSVAGRVSRREKADAKLGNLSLKLKEGDRLD 1620
 DB 1561 TCMPGRTLYLLASPDPKQKRWATLALSVAGRVSRREKADAKLGNLSLKLKEGDRLD 1620
 QY 1621 MNCITLPSDDVNLVGRTEGILYALNVLKNSLTHVPGIAGVQIYIIXDLEKLMJAGERA 1680
 DB 1621 MNCITLPSDDVNLVGRTEGILYALNVLKNSLTHVPGIAGVQIYIIXDLEKLMJAGERA 1680
 QY 1681 LCLVDYKVKQSLAQSILPAQDPISPNIFEAVNGCHLFGAGKIENTGLCIQAMPKSVIIL 1740
 DB 1681 LCLVDYKVKQSLAQSILPAQDPISPNIFEAVNGCHLFGAGKIENTGLCIQAMPKSVIIL 1740
 QY 1741 RYNNENISKYCIKRIETSEPCSCIHFTNYSILIGTNKFYEIDMKQYTLBEFLDKNDHSLA 1800
 DB 1741 RYNNENISKYCIKRIETSEPCSCIHFTNYSILIGTNKFYEIDMKQYTLBEFLDKNDHSLA 1800
 QY 1801 PAYPAASNSFPVSIYQVNSAGQREBYLCEHFGVFDVSGRRSRDILKMSRLPLAFA 1860
 DB 1801 PAYPAASNSFPVSIYQVNSAGQREBYLCEHFGVFDVSGRRSRDILKMSRLPLAFA 1860
 QY 1861 YREBYLFTVHFNLSLEVEIQARSSAGTPARAYLDIPNRYLGPALSSGAIYLAASSYODKL 1920
 DB 1861 YREBYLFTVHFNLSLEVEIQARSSAGTPARAYLDIPNRYLGPALSSGAIYLAASSYODKL 1920
 QY 1921 RVICCKNLVKSQTEHHRGPSTSRSSPNKRGPTTYNEHITKRVASSPADPEGSHPREP 1980
 DB 1921 RVICCKNLVKSQTEHHRGPSTSRSSPNKRGPTTYNEHITKRVASSPADPEGSHPREP 1980
 QY 1981 STPHRYAEGTELRDQSPERPLEREKSPGRITSTRERSPARLPEDSSRGRLPAGAVRT 2040
 DB 1981 STPHRYAEGTELRDQSPERPLEREKSPGRITSTRERSPARLPEDSSRGRLPAGAVRT 2040
 QY 2041 PLSQVNVKMDQSSV 2054
 DB 2041 PLSQVNVKMDQSSV 2054

RESULT 2
 US-10-028-946-4
 ; Sequence 4, Application US/10028946
 ; Patent No. 6734009
 ; GENERAL INFORMATION:
 ; APPLICANT: Yu, Xuanchuan
 ; APPLICANT: Miranda, Maricar
 ; APPLICANT: Fridde, Carl Johan
 ; TITLE OF INVENTION: No. 6734009e1 Human Kinases and Polynucleotides Encoding the Same
 ; FILE REFERENCE: LEX-0289-USA
 ; CURRENT APPLICATION NUMBER: US/10/028,946
 ; PRIOR FILING DATE: 2001-12-20
 ; PRIOR APPLICATION NUMBER: US 60/258,335
 ; NUMBER OF SEQ ID NOS: 4
 ; SOFTWARE: PatsSeq for Windows Version 4.0
 ; SEQ ID NO 4
 ; LENGTH: 1958
 ; TYPE: PRT
 ; ORGANISM: homo sapiens
 US-10-028-946-4

Query Match 94.5%; Score 9916; DB 4; Length 1958;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1945; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MUKFKYGARNPLDGAAPLARSRLNLFQCKAPPMTOQNSPLSRBGLIDLALFVLFE 60
 DB 1 MUKFKYGARNPLDGAAPLARSRLNLFQCKAPPMTOQNSPLSRBGLIDLALFVLFE 60

TYPE: PRT
ORGANISM: Human
US-09-804-471A-2

Query Match 23.3%; Score 2440; DB 4; Length 497;
Best Local Similarity 99.6%; Pred. No. 1.1e-118;
Matches 466; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLKFKYGARNPLDGAAPLPIASRASRLNLFPGQKPPMTQOQMSPLSREGILDALFVLFPE 60
DB 1 MLKFKYGARNPLDGAAPLPIANRASRLNLFPGQKPPMTQOQMSPLSREGILDALFVLFPE 60
QY 61 ECGSPALMKIKHVSNFPRKYSDDTIAELQELQPSAKDFEVSILVCGGFAEVQVVRKATG 120
DB 61 ECGSPALMKIKHVSNFPRKYSDDTIAELQELQPSAKDFEVSILVCGGFAEVQVVRKATG 120
QY 121 DIYAMKVMKKKALLAOQOVSPFEEERNILSRSTSPMIPOQYAFODKNHLYLWMEYQPGG 180
DB 121 DIYAMKVMKKKALLAOQOVSPFEEERNILSRSTSPMIPOQYAFODKNHLYLWMEYQPGG 180
QY 181 DLISLNRVEDQDENLQFYLAELILAVSHVLMGVHRDIRENILVDRTGHIKLVDF 240
DB 181 DLISLNRVEDQDENLQFYLAELILAVSHVLMGVHRDIRENILVDRTGHIKLVDF 240
QY 241 GSAAKNSNMVNAKLPIGTPDYMAPEVLTVMNGDGKTYGLDCDMSVGVIAEMYYGR 300
DB 241 GSAAKNSNMVNAKLPIGTPDYMAPEVLTVMNGDGKTYGLDCDMSVGVIAEMYYGR 300
QY 301 SPFAEGTSARTFNINNFQFLKFPDDPKVSDPDLIOSLLCGQKRLKFEGLCCHPFF 360
DB 301 SPFAEGTSARTFNINNFQFLKFPDDPKVSDPDLIOSLLCGQKRLKFEGLCCHPFF 360
QY 361 SKIDMNNIRNSPPFPVTLKSDDDTSNPFDEPKNSWSSPCQSPSGFSGEELPFVGF 420
DB 361 SKIDMNNIRNSPPFPVTLKSDDDTSNPFDEPKNSWSSPCQSPSGFSGEELPFVGF 420
QY 421 YSKALGILGRSESVSGLDSPAKTSMKKLLIKSKELQDSQDKCHKV 468
DB 421 YSKALGILGRSESVSGLDSPAKTSMKKLLIKSKELQDSQDKCHKV 468

RESULT 4
US-10-238-709-2
Sequence 2, Application US/10238709
Patent No. 6680188
GENERAL INFORMATION:
APPLICANT: WEBSTER, Marion et al
TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
FILE REFERENCE: C0001164DI
CURRENT APPLICATION NUMBER: US/10/238,709
CURRENT FILING DATE: 2002-09-11
NUMBER OF SEQ ID NOS: 4
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 2
LENGTH: 497
TYPE: PRT
ORGANISM: Human
US-10-238-709-2

Query Match 23.3%; Score 2440; DB 4; Length 497;
Best Local Similarity 99.6%; Pred. No. 1.1e-118;
Matches 466; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLKFKYGARNPLDGAAPLPIASRASRLNLFPGQKPPMTQOQMSPLSREGILDALFVLFPE 60
DB 1 MLKFKYGARNPLDGAAPLPIANRASRLNLFPGQKPPMTQOQMSPLSREGILDALFVLFPE 60
QY 61 ECGSPALMKIKHVSNFPRKYSDDTIAELQELQPSAKDFEVSILVCGGFAEVQVVRKATG 120
DB 61 ECGSPALMKIKHVSNFPRKYSDDTIAELQELQPSAKDFEVSILVCGGFAEVQVVRKATG 120

QY 121 DIYAMKVMKKKALLAOQOVSPFEEERNILSRSTSPMIPOQYAFODKNHLYLWMEYQPGG 180
DB 121 DIYAMKVMKKKALLAOQOVSPFEEERNILSRSTSPMIPOQYAFODKNHLYLWMEYQPGG 180
QY 181 DLISLNRVEDQDENLQFYLAELILAVSHVLMGVHRDIRENILVDRTGHIKLVDF 240
DB 181 DLISLNRVEDQDENLQFYLAELILAVSHVLMGVHRDIRENILVDRTGHIKLVDF 240
QY 241 GSAAKNSNMVNAKLPIGTPDYMAPEVLTVMNGDGKTYGLDCDMSVGVIAEMYYGR 300
DB 241 GSAAKNSNMVNAKLPIGTPDYMAPEVLTVMNGDGKTYGLDCDMSVGVIAEMYYGR 300
QY 301 SPFAEGTSARTFNINNFQFLKFPDDPKVSDPDLIOSLLCGQKRLKFEGLCCHPFF 360
DB 301 SPFAEGTSARTFNINNFQFLKFPDDPKVSDPDLIOSLLCGQKRLKFEGLCCHPFF 360
QY 361 SKIDMNNIRNSPPFPVTLKSDDDTSNPFDEPKNSWSSPCQSPSGFSGEELPFVGF 420
DB 361 SKIDMNNIRNSPPFPVTLKSDDDTSNPFDEPKNSWSSPCQSPSGFSGEELPFVGF 420
QY 421 YSKALGILGRSESVSGLDSPAKTSMKKLLIKSKELQDSQDKCHKV 468
DB 421 YSKALGILGRSESVSGLDSPAKTSMKKLLIKSKELQDSQDKCHKV 468

RESULT 5
US-09-804-471A-4
Sequence 4, Application US/09804471A
Patent No. 6479269
GENERAL INFORMATION:
APPLICANT: WEBSTER, Marion et al
TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
FILE REFERENCE: C0001164
CURRENT APPLICATION NUMBER: US/09/804,471A
CURRENT FILING DATE: 2001-03-13
NUMBER OF SEQ ID NOS: 4
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 4
LENGTH: 494
TYPE: PRT
ORGANISM: Mus Musculus
US-09-804-471A-4

Query Match 20.7%; Score 2175.5; DB 4; Length 494;
Best Local Similarity 88.2%; Pred. No. 5.2e-105;
Matches 413; Conservative 25; Mismatches 29; Indels 1; Gaps 1;

QY 1 MLKFKYGARNPLDGAAPLPIASRASRLNLFPGQKPPMTQOQMSPLSREGILDALFVLFPE 60
DB 1 MLKFKYGARNPLDGAAPLPIASRASRLNLFPGQKPPMTQOQMSPLSREGILDALFVLFPE 60
QY 61 ECGSPALMKIKHVSNFPRKYSDDTIAELQELQPSAKDFEVSILVCGGFAEVQVVRKATG 120
DB 61 ECGSPALMKIKHVSNFPRKYSDDTIAELQELQPSAKDFEVSILVCGGFAEVQVVRKATG 120
QY 121 DIYAMKVMKKKALLAOQOVSPFEEERNILSRSTSPMIPOQYAFODKNHLYLWMEYQPGG 180
DB 121 DIYAMKVMKKKALLAOQOVSPFEEERNILSRSTSPMIPOQYAFODKNHLYLWMEYQPGG 180
QY 181 DLISLNRVEDQDENLQFYLAELILAVSHVLMGVHRDIRENILVDRTGHIKLVDF 240
DB 181 DLISLNRVEDQDENLQFYLAELILAVSHVLMGVHRDIRENILVDRTGHIKLVDF 240
QY 241 GSAAKNSNMVNAKLPIGTPDYMAPEVLTVMNGDGKTYGLDCDMSVGVIAEMYYGR 300
DB 241 GSAAKNSNMVNAKLPIGTPDYMAPEVLTVMNGDGKTYGLDCDMSVGVIAEMYYGR 300
QY 301 SPFAEGTSARTFNINNFQFLKFPDDPKVSDPDLIOSLLCGQKRLKFEGLCCHPFF 360
DB 301 SPFAEGTSARTFNINNFQFLKFPDDPKVSDPDLIOSLLCGQKRLKFEGLCCHPFF 360

QY 361 SKIDMNNIRNSPPFVPTLKSDDDTSNFDEPKNSWSSPCQLSPSGSGEELPFVGF 420
 Db 360 ARTMNNIRNSPPFVPTLKSDDDTSNFDEPKNSWAFILCVPAEPLAFSGEELPFVGF 419
 QY 421 YSKALGILGRSESVSGLDSPAKTSMEKLLIKSKELQSDCKHKM 468
 Db 420 YSKALGILGRSESVSSLDSPAKYSMEKLLIKSKELQSDCKCHKV 467

RESULT 6

US-10-238-709-4
 ; Sequence 4, Application US/10238709
 ; Patent No. 6680188
 ; GENERAL INFORMATION:
 ; APPLICANT: WEBSTER, Marion et al
 ; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
 ; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
 ; FILE REFERENCE: C001164D1V
 ; CURRENT APPLICATION NUMBER: US/10/238,709
 ; CURRENT FILING DATE: 2002-09-11
 ; NUMBER OF SEQ ID NOS: 4
 ; SOFTWARE: FASTSEQ for Windows Version 4.0
 ; SEQ ID NO 4
 ; LENGTH: 494
 ; TYPE: PRT
 ; ORGANISM: Mus Musculus
 US-10-238-709-4

Query Match 20.7%; Score 2175.5; DB 4; Length 494;

Best Local Similarity 88.2%; Pred. No. 5.2e-105; Mismatches 25; Indels 1; Gaps 1;

QY 1 MLKFKYGARNPLDGAAPLIRASRLNLFQGGKPPMTQOQMSPLSREGILDLFVLFE 60
 Db 1 MLKFKYGARNPEASAPLIRASRLNLFQGGKPPMTQOQMSPLSREGILDLFVLFE 60
 QY 61 ECSOPALMKIKHVSNFVFKYSDTIAELOLQPSAKDFEVSIVCGHFAEYQVVRKATG 120
 Db 61 ECSOPALMKIKHVSFVQKYSDTIAELRELQPSARDFEVSIVCGHFAEYQVVRKATG 120
 QY 121 DIYAMKVMKKALLAQOVSFEEERNILSRSTSPWIPQIQYAFQDKNHLVLYMEYOPGG 180
 Db 121 DIYAMKVMKKALLAQOVSFEEERNILSRSTSPWIPQIQYAFQDKNHLVLYMEYOPGG 180
 QY 181 DLISLARYEDQDENLIOFYLAELILAVSHVHMGVYHRDIKENLILVDTGHIKLVDF 240
 Db 181 DLISLARYEDQDENLIOFYLAELILAVSHVHMGVYHRDIKENLILVDTGHIKLVDF 240
 QY 241 GSAAKNNSKNVNAKLPIGTEDYNAPEVLTVMNGDGKGTGGLDCDMSVGVAYEMITGR 300
 Db 241 GSAAKNNSKN-VDAKLPIGTEDYNAPEVLTVMNEDRGITGGLDCDMSVGVAYEMITGR 299
 QY 301 SPFAEGTISARTENNINMFORFLKFPDDPKVSSDFLDLIQSLICGQKELKEGICHPFF 360
 Db 300 TPFEETGTSATFNINMFORFLKFPDDPKVSSSELDLQSLICVQKRLKEGICHPFF 359
 QY 361 SKIDMNNIRNSPPFVPTLKSDDDTSNFDEPKNSWSSPCQLSPSGSGEELPFVGF 420
 Db 360 ARTMNNIRNSPPFVPTLKSDDDTSNFDEPKNSWAFILCVPAEPLAFSGEELPFVGF 419
 QY 421 YSKALGILGRSESVSGLDSPAKTSMEKLLIKSKELQSDCKHKM 468
 Db 420 YSKALGILGRSESVSSLDSPAKYSMEKLLIKSKELQSDCKCHKV 467

RESULT 7

US-09-916-204-2
 ; Sequence 2, Application US/09916204
 ; Patent No. 6638745
 ; GENERAL INFORMATION:
 ; APPLICANT: WEI, Ming-Hui et al.
 ; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC

; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
 ; FILE REFERENCE: C001164C1P
 ; CURRENT APPLICATION NUMBER: US/09/916,204
 ; CURRENT FILING DATE: 2001-07-24
 ; NUMBER OF SEQ ID NOS: 6
 ; SOFTWARE: FASTSEQ for Windows Version 4.0
 ; SEQ ID NO 2
 ; LENGTH: 257
 ; TYPE: PRT
 ; ORGANISM: Human
 US-09-916-204-2

Query Match 12.3%; Score 1287; DB 4; Length 257;

Best Local Similarity 100.0%; Pred. No. 2.1e-59; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLKFKYGARNPLDGAAPLIRASRLNLFQGGKPPMTQOQMSPLSREGILDLFVLFE 60
 Db 1 MLKFKYGARNPLDGAAPLIRASRLNLFQGGKPPMTQOQMSPLSREGILDLFVLFE 60
 QY 61 ECSOPALMKIKHVSNFVFKYSDTIAELOLQPSAKDFEVSIVCGHFAEYQVVRKATG 120
 Db 61 ECSOPALMKIKHVSNFVFKYSDTIAELOLQPSAKDFEVSIVCGHFAEYQVVRKATG 120
 QY 121 DIYAMKVMKKALLAQOVSFEEERNILSRSTSPWIPQIQYAFQDKNHLVLYMEYOPGG 180
 Db 121 DIYAMKVMKKALLAQOVSFEEERNILSRSTSPWIPQIQYAFQDKNHLVLYMEYOPGG 180
 QY 181 DLISLARYEDQDENLIOFYLAELILAVSHVHMGVYHRDIKENLILVDTGHIKLVDF 240
 Db 181 DLISLARYEDQDENLIOFYLAELILAVSHVHMGVYHRDIKENLILVDTGHIKLVDF 240
 QY 241 GSAAKNNSKNV 252
 Db 241 GSAAKNNSKNV 252

RESULT 8

US-10-282-048-2
 ; Sequence 2, Application US/10282048
 ; Patent No. 6692948
 ; GENERAL INFORMATION:
 ; APPLICANT: WEI, Ming-Hui et al.
 ; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
 ; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
 ; FILE REFERENCE: C001164C1P-DIV
 ; CURRENT APPLICATION NUMBER: US/10/282,048
 ; CURRENT FILING DATE: 2002-10-29
 ; NUMBER OF SEQ ID NOS: 6
 ; SOFTWARE: FASTSEQ for Windows Version 4.0
 ; SEQ ID NO 2
 ; LENGTH: 257
 ; TYPE: PRT
 ; ORGANISM: Human
 US-10-282-048-2

Query Match 12.3%; Score 1287; DB 4; Length 257;

Best Local Similarity 100.0%; Pred. No. 2.1e-59; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLKFKYGARNPLDGAAPLIRASRLNLFQGGKPPMTQOQMSPLSREGILDLFVLFE 60
 Db 1 MLKFKYGARNPLDGAAPLIRASRLNLFQGGKPPMTQOQMSPLSREGILDLFVLFE 60
 QY 61 ECSOPALMKIKHVSNFVFKYSDTIAELOLQPSAKDFEVSIVCGHFAEYQVVRKATG 120
 Db 61 ECSOPALMKIKHVSNFVFKYSDTIAELOLQPSAKDFEVSIVCGHFAEYQVVRKATG 120
 QY 121 DIYAMKVMKKALLAQOVSFEEERNILSRSTSPWIPQIQYAFQDKNHLVLYMEYOPGG 180
 Db 121 DIYAMKVMKKALLAQOVSFEEERNILSRSTSPWIPQIQYAFQDKNHLVLYMEYOPGG 180

QY 181 DLISLNRVYEDQDENTLIQFLAELILAVSHLMGYVHDIKRPENILVDRTHGIKLVDF 240
 Db 181 DLISLNRVYEDQDENTLIQFLAELILAVSHLMGYVHDIKRPENILVDRTHGIKLVDF 240
 QY 241 GSAKNNNNKVV 252
 Db 241 GSAKNNNNKVV 252

RESULT 9
 US-08-685-576-1
 ; Sequence 1, Application US/08685576
 ; Patent No. 5906819
 ; GENERAL INFORMATION:
 ; APPLICANT: Kaibuchi, Kozo
 ; APPLICANT: Imamatsu, Akihiro
 ; APPLICANT: Nakano, Takeshi
 ; APPLICANT: Ito, Masaaki
 ; APPLICANT: Takahashi, No. 5906819uaki
 ; TITLE OF INVENTION: RHO TARGET PROTEIN RHO-KINASE
 ; NUMBER OF SEQUENCES: 16
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Foley & Lardner
 ; STREET: 3000 K Street, N.W., Suite 500
 ; CITY: Washington
 ; STATE: D.C.
 ; COUNTRY: USA
 ; ZIP: 20007-5109
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patentin Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/685,576
 ; FILING DATE: 24-JUL-1996
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: JP 7-325129
 ; FILING DATE: 20-NOV-1995
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: JP 8-17150
 ; FILING DATE: 05-JAN-1996
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: JP 8-131206
 ; FILING DATE: 26-APR-1996
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Bent, Stephen A.
 ; REGISTRATION NUMBER: 29,768
 ; REFERENCE/DOCKET NUMBER: 16887/843
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (202) 672-5300
 ; TELEFAX: (202) 672-5399
 ; TELEX: 904136
 ; INFORMATION FOR SEQ ID NO: 1:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1388 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-08-685-576-1

Query March 12.0%; Score 1263.5; DB 2; Length 1388;
 Best local similarity 24.4%; Pred. No. 2.1e-57;
 Matches 382; Conservative 274; Mismatches 466; Indels 443; Gaps 41;

QY 15 GAAPLIA---SRASRLNLFQGGKPPFTQOQMSPLSREGILDALFVLFEECSGPALMKIK 71
 Db 12 GAAPEVSGDGAGASR-----QRKLEALIRDRSPINIVSSLDGLNPLVLDLDFPALRKNK 66
 QY 72 HVSFVFKYSTTIALDELQPSADQFVRSVVGCGHFAEVQVVEKATGDIYANKWKKK 131
 Db 72 HVSFVFKYSTTIALDELQPSADQFVRSVVGCGHFAEVQVVEKATGDIYANKWKKK 131

Db 67 NIDNPLNRYEKIYKIRGLQMKABDYDVVYVIGRGAFGVEQLVBRHKASQKVYAMKLSKF 126
 QY 132 ALIAQEVSPFEEERNILSRSTPMIPOLQYAFQDKNHLVLYVNEYQGGLLSLNRYED 191
 Db 127 EMIKRSDSAFEEERDIMAFAVNSPWWQLFCAFOQDKYLVVMEYMGGLVNLMSYD- 185
 QY 192 QLDENLIQFLAELILAVSHLMGYVHDIKRPENILVDRTHGIKLVDFGSAKNNNNK 251
 Db 186 -VEPKMAKFTYALVNLADLHSMGLHSDKFLHDKFNNLMDLHKGHLKLADFCCKMDSTGM 244
 QY 252 VNAKLPIGTPDYNAPEVLTVMNGDGKGYGLDQDMSVGVIAVEMIYGRSPFAEGTSART 311
 Db 245 VHCDTAVGTPDYISPEVLSKQGGD--GYGRBCDMSVGVFLPEMLVGPDPFADSLVGT 302
 QY 312 FNNINAFQRLKPPDPKYSDFLDLISLCSQKRLKEG--LCHEPFESKIDWN-- 366
 Db 303 YSKIMDKHNSLCPEPDAEISKHAKNLCVLTREYVLEGNVGEELKQHPFNQDQNMWD 362
 QY 367 NINNSPPFPVPTLKSDDDNSNPEPEKNSW-VSSPQQLSPSGSGELPFGVSGSKAL 425
 Db 363 NIRETAPVPELSDDISNFDIEDDKGDEVTFPI--PKAFVGNQLPFIQFTYREN 419
 QY 426 GILGRSESVSGDSPA--KTSMEKKLISKXELQDSQDKCHMEQEMTRLHRRVSEVE 483
 Db 420 LILS-----DSPCKENDSIQSR--KNEBSQEIQKKVLTLEHLS----- 457
 QY 484 AVLSQKEVEIKASETQSLIEQDLATYITECSSLKSLEQARNEVSOEDKALQLHDIR 543
 Db 458 -----TEIQAKER-----LEQ-----KCKSVNTRLEKAVAKELEBEITLRKRVSETLR 499
 QY 544 --EQSRKLOIKEOYQAOVEEMRLMMNQLEELVBARRRSDIYESLRSERLAAEFKR 601
 Db 500 QLRERKALQHKAEYQKADHEADKRNLENVSNLSKQLEJLKKRNNQSISTE----- 555
 QY 602 KATECOHKLKAKDOGKPEVEYAKLEKINAEOQLKIOELO-----X 644
 Db 556 KVNQLOQLDETALARTESDTAARLKQABESSKQIQLESNNRDLQDKNCLEETAKK 615
 QY 645 LEKAVASTATLONIROAKERAERELEKLN-----EDSSG-IRKKLVAEERRH 698
 Db 616 LEKEF--INLQSVLESERDRTHGEIINDLGRIISGLEEDVYNGKILAKTELEKROL 672
 QY 699 SLENKVRLTEWRRERNRLKDDIQTSQOIQOMADKILEEKHREARQVSAQHLEVTLKQ 758
 Db 673 Q-----ERFTDLKEKNMEIDMTYQKVIQOS-----LEQEEYENKA-----TKARLAD 717
 QY 759 KEQHYE-----EKIKVLDNQIKDLADKETLENMQRHE-----BEAHEKGI 801
 Db 718 KKKIYESIEAKSEAMKEMKKLSERTLKQYENLLBAEKRCSTLIDCLKSQOKINE 777
 QY 802 LSFQKAMINAMSKIRSLSEQRIYELSEANKLAANSLFTQRMKKAQEMISLRQOKFYL 861
 Db 778 LTKQKQVTLN--EDVNLTLKIEQETQKCL-----TQNDLKMQTQOVVNTLMS----- 823
 QY 862 ETQAGKLEAQRNRLLEQLEKISHQDSDKYRLLETRLRLEVSLEHBEOKELKROUTEL 921
 Db 824 -----EQQLKQ-----ENNHLLEMKQSL-----EKQVAELRKEQDA 855
 QY 922 QLSLOERESQLTALQARAALLESQLRQAKTELEETT--AAEAEIQLATNHRDEIQRKF 978
 Db 856 DQGMKELODLEBAEQFSTYIKQVAREIKERCEKTKLCELELQKKQLOLDERDLSLAQL 915
 QY 979 DALRNSCTVTIDLEBOLNQITEDNAELNNQFYLSQLEDAAGANDIYQLRSEVDHLRR 1038
 Db 916 E-----ITLRKADSEQLARSIAAEQVSDLEKEKIMKE-----LEIKEMMARHKQ 959
 QY 1039 EITEREMQLTSLQQTMEALKTTCTMLEBOVMDEALINDELLEKERQWEARSVLGDE-- 1095
 Db 960 ELTEKQATIASLETRTTLTSDVANLAN--EKKEELNNKIKEAQEO--LSLRKQBEIS 1012
 QY 1096 -----KSQFECRVNELQMDLTERKQSPARADQRTESQVAVLAVKEHKEILALQALKE 1151
 Db 1013 AALIKAOFE-----KQLTLTE 1027

QY 1152 OKLAEISLTDKLNDEKKHMLMNNARSLOQKLETERELKORLLEQAKLOOQNDLOKNH 1211
Db 1028 RTLTKQAVNKLAETLNKKEPVKRGNDTVRRKKEKNRLHMLKSEKREKLTQOM----- 1081
QY 1212 IFRULQGLQALDRADILKTERSDLEYQLENIQVLYSHEKVMESTGTSQTKLIDFLQAK 1271
Db 1082 -----IKYQELNEMQALBESQI----- 1101
QY 1272 MDQAKKKKVPLOYNELKALKEKARCALBESALQKTRIELRSAREBAHRKATDHPHP 1331
Db 1102 -----RLQMTLDSKDSQIEQRLSQALHIGLDS----- 1133
QY 1332 STPATARQOIMSAIVRSPEHQPSAMSLAPSSSRKSSSTPEESRRLKRMHNTPHR 1391
Db 1134 -----SIGSGPGTDEADGFPES----- 1151
QY 1392 FNVGLNMRATKCAVCLDTVHFGRQASKLEQVCMHPCSTCLPATGSLPAEYATHPTFA 1451
Db 1152 ----- 1151
QY 1452 PCRDKMNSPGLOKTEPSSSLHLEGMMKVPRNNKSGOQGMKRYIVLESGKVLIDNEARE 1511
Db 1152 -----RLBGLSLPVANNTKKFGWVKYIVVSKLILFYDSEQDK 1191
QY 1512 AGQRP 1516
Db 1192 EQSNP 1196

RESULT 10
US-08-685-871-2
Sequence 2, Application US/08685871
Patent No. 6013499
GENERAL INFORMATION:
APPLICANT: NARUMIYA, Shuh
APPLICANT: IWAMATSU, Akihiro
TITLE OF INVENTION: RHO TARGET PROTEIN KINASE P160
NUMBER OF SEQUENCES: 68
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W., Suite 500
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/685,871
FILING DATE: 24-JUL-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 8-184102
FILING DATE: 25-JUN-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 7-262553
FILING DATE: 14-SEP-1995
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 16887/845
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 672-5300
TELEFAX: (202) 672-5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1354 amino acids
TYPE: amino acid

TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-685-871-2
Query Match 12.0%; Score 1260; DB 3; Length 1354;
Best Local Similarity 26.3%; Pred. No. 3,26-57;
Matches 400; Conservative 293; Mismatches 531; Indels 296; Gaps 53;
QY 44 SPLRSIGILDALFVLFEBCSQPALMKIKHVSNFVRKYSDTIAELQELQPSAKDEPVSLV 103
Db 23 SEVNSDCLLDGLDALVVDLPDPALRKNNKINDFLSRYKDTINKIRDLRMKAEDEVVAVKI 82
QY 104 GCGHFAVQVVRKATGTDIYAMKMKKALLAQOVSPFEEBRNIIILSRSTPMLPOLOYA 163
Db 83 GRGAFGEVOLVRNHSSTKRYVAMKLLSKFEMIKRSDSAFWEERIDIMAFANSPWVQLFYA 142
QY 164 FQDKNHLVLMVEYOGGDLILNRYDQDLENLIGFVLAELIAYNSVHLMGVVHRDIK 223
Db 143 FQDRYILMVEIYFPGDLVNLMSNYD--VPEKARFTYLAELVLDALHISMGFIHRDVK 200
QY 224 PENILVDRGTGHIKLVDEGSAKXNSNINNVNAKLPITGPDYMAPEVLTVMNGDGKTYGLD 283
Db 201 PDNMLDKSGHGLKADHGTGCMKNKKBGMVACDTAVGTPDYISPEVLKSQGD--GYGRE 258
QY 284 CDWWSGVIAVEMTYGSPPAEGTSARTENNINMFOQLFKPPDPKVSDFLLIQSLLC 343
Db 259 CDWWSGVFLYEMLVGDTPEFADSLVGTYSKIMHKKSLTFPPDNDISKEAKNLICAFLT 318
QY 344 GQKERLKFEG---LCCHPFSKID--WNINRSPRPVPTLKSDDPDSNDEPERKSNVS 398
Db 319 DREYLRGRNGVEELKRLFFKNDQWAEETLDYAPVAPVPLSSDIDTSNDDLEEDGEE 378
QY 399 SS-PCQLSPSGFSGEELPFVGFYSYKALGILGSESVSGIDSPAKTSMEKULLIKSKE 457
Db 379 ETFFI---PKAFVGNQLPVGFY-----YSNRYSANPNDRNRTSSN-----ADKS 423
QY 458 LQDS-QDKCHRMGEQETRLHRSVSEVAIVSQEVELKASQETOR--SLBQD-LATYIT 512
Db 424 LQESLQKTIYKLEBQ-----HNEMQLKDEMEQKCRYSNIKLDIKMKELD 468
QY 513 ECSSLKSLQANNEVGQEDDKALQLLHDIREQSKIQELKEQYQOVE-EMKLMNQ 571
Db 469 BEGNQRNLEST--VSQIKERKLLQHRINEYQKRAQ--ENKRNVENEVSTLQDL 523
QY 572 EEDLVASRRSDIYSESLRESRLAEBFKRATKTECOHKLAKXQGRVEGEVYALEKIN 631
Db 524 EDUKKYSQNSQL-----AND--KLSQLOQLSEANDLRTESDTAVRLKSH 568
QY 632 AEOQLKIQLQELKELKAVKASTATELLONIROAKERARELEKQNRDSSSEGRKKLV 691
Db 569 TEMSKISQI--ESLNRELQ--ERNRILEN--SKSQTDQVYQLQ-----AIL 610
QY 692 EAEBRRSLKMKYKRLTETMERREKRLKODIQTK---SQOIQOMADKILEEKKHRAQV 747
Db 611 EAERDRG-----HDSMIGDQARTISLQEBVKIKHLKELVEGGRKKAQD 657
QY 748 SAQHLVHLKQKHVEKIKVLDNOIKQDLADETKELNMOHREBAHE--KGIILS 803
Db 658 MLANSE---KEK-----NLNLEIDLNTK--IKSLQQRLEQGVNHEHKTAKRLTD 700
QY 804 EOKAMINAMDSKIRSEORIVELSEANKLANSLFTORMKAOQEMIS-ELRQOQFYLE 862
Db 701 KHQISIEAKSVANCEMKKLEBERAREKENVV---QIEKQCSMLVDVLKQSQKLE 756
QY 863 TQNGKLEAQRKLEBQLEKISHQDSDKRLLELFTRLREVSLSJBHEQK----- 911
Db 757 HLTGNK-----RMEDVKNTLTLQEBQSNRLLQLQNELKTAQAFADLTKLEKQMQKEIN 812
QY 912 --LELKRQLTELQSL-----LOERESQLTALQAAABAASQQLQATLEBETT 957
Db 813 TLLEAKR-LLEFELAQTIKYRGNEGOMRELQQLLEBOYFSTLYKTIQVELKEEIBKN 871
QY 958 ABAEETIQALTARDEIQRFDALRNSCTVITDLBQLNQLTEBDNMLNQNFFYLSKQD 1017


```

Db      679 LEKESNMEDMTYQVKIQQS-----LEGEAEHKA-----TKARLADKKKIYESIEEA 728
Qy      765 --EKIYVLDNQIKKOLADKETLENMORHE-----BEAEKGIUSEQAMINAM 812
Db      729 KSEAMKEMEKELBERTLKQVENILLAEKRGSLDDCDLQSQQKINELLKQDVAN-- 786
Qy      813 DSKRISLEQRIVEISEANKLAANSLSLTFQRMKAQOEMISELRQOKFYLETQAGKLEAQN 872
Db      787 -EDVRNLTLKIEQETQKCL-----TQNDLKMQTOQVNTLKMS----- 823
Qy      873 RKSEOLEKISHQSHSDKNRLLELETRELSLHEBQKLEKQKQLELOLSQRESQ 932
Db      824 --EKQKQ-----ENHLMEMKMNL-----EKQNAELRKERQADQOMKELDQOL 866
Qy      933 TALQADAALESQURQAKTELEETT--AEAEERIQALTARDEIQKFDALRNSCVIT 989
Db      867 EAEYFSTVYQVRELKEBCEKTKQKELQKQKQKQLEODERDLSLAQLE-----IT 918
Qy      990 DLEBQNLQLEEDNAELNNQNFYLSKQLEASGANDVQLRSEVDHLRREITEREMOLTS 1049
Db      919 LTKADSQELASISIEEYOYSDLEKEKIME-----LEIEMMARHKOELTEKDATIAS 970
Qy      1050 QKQTMELKTCMLEBQVMDLEALNDELLEKERQWEMBSVLGE-----KSOFECR 1102
Db      971 LEBNRTLSVANIAM--EKELNMLKQVQOQ--LSRLQDEISAAITQOPE-- 1021
Qy      1103 VRELQMLDTEKQSRARADQRTESRQVVELAVENKAKAILDALQALKEQKLKAKESLSDK 1162
Db      1022 -----KQLTERTLKQAVNKL 1038
Qy      1163 LNDLEKKHAMELMMNARSLOQKLETERELKQRLLEBQALQOQMDLQKNNIFRLTQGLEA 1222
Db      1039 AEIMNRKEPVKRGNDIVRKRKEKRNKLHMKELKEREGLTQOM----- 1081
Qy      1223 LDRADLKTERSDELYOLENIQVLYSHKVMMEGTISQOTKLIFPLQAKMDQPAKKKKVP 1282
Db      1082 -----IKYOKEMLNMQOIAEESQI----- 1101
Qy      1283 LQVNELKALEKEKARCABELEALQKTRIELRSAREBAHAKATDHPHPTPATARQOIA 1342
Db      1102 --RELQMTLDSKQSDLEBQLSQALAHIGLDS-- 1133
Qy      1343 MSALVRSPEHQPSAMSLAPSSSRKESSTPEEFSRRLKERMHNITPHRFVNGIAMBATK 1402
Db      1134 -----SIGSGPGDAEADGPFES----- 1151
Qy      1403 CAVCLDTVHFGROASKLEECQVMCHPKCSTCLPATCGLPAEYATHTFEACGRDMNSPGL 1462
Db      1152 ----- 1151
Qy      1463 QTEPSSSLHEGMKVPRNNKRGQOGMDRKVYLVEGSKVLYDNEAREAGORP 1516
Db      1152 -----RLEGWLSLPRVKNNTKKGFWKCVIVSSKILFYDSEODKQSNP 1196

RESULT 12
US-09-976-594-296
; Sequence 296, Application US/09976594
; Patent No. 6673549
; GENERAL INFORMATION:
; APPLICANT: Furness, Michael
; APPLICANT: Buchbinder, Jenny
; TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELL CULTURES TREATED WITH STEROIDS
; FILE REFERENCE: PA-0041 US
; CURRENT APPLICATION NUMBER: US/09/976,594
; PRIOR FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: 60/240,409
; NUMBER OF SEQ ID NOS: 1143
; SOFTWARE: PERL Program
; SEQ ID NO 296
; LENGTH: 1388

```

```

; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6673549 064987CD1
US-09-976-594-296

Query Match      12.0%; Score 1257; DB 4; Length 1388;
Best Local Similarity 24.6%; Pred. No. 4,7e-57;
Matches 375; Conservative 279; Mismatches 485; Indels 388; Gaps 37;

Qy      24 ASRLNLFQOKKPRPMTOQOMSPLSREGILDALFVLFECSCPALMKIKHVSNFVRKYSDT 83
Db      24 ASR-----QKLEALINDPSPSPINVESLLGLNSLVLDLPFALPKKNNKINDFNRIEKT 78
Qy      84 IAELOEPQAKQFEVSLVGCGFHFAEVQVVRREKATGDIYAMKVMKKKALLAQOVSFFE 143
Db      79 VKIRIGIQMAEDVDYKVGIRGAFGEVQVLRHNASQKVYAMKLSIFEMIKRDSAFW 138
Qy      144 EERNILSRSTSPWIPOLQVAFQDKNHLVYVNEVOPGDDLILNRYEDQDENLIOFYLA 203
Db      139 EERDIMAFAANSPMVVQLFYAFQDDRILYVMEVYVPGDLVNLMSNYD--VPEKAKFYTA 196
Qy      204 ELIAVSHVLMGVNHDIPENILVDRTHIKLVDSGAAXNNNSQVNAKLPITGPDY 263
Db      197 EVVIALDAISHMGLIHRDVPRDMLDKHGLKLDAGTCMKMDGTGMVHCDTAVAGSPDY 256
Qy      264 MAPEVLTVNNAGDGKGTGLCDMWSVGVIAYEMIVGSRPRAEGSARTFNNINMFORFLK 323
Db      257 ISPEVLKSGQGD--GFYGRCDMWSVGVIYEMIVGTPFYRADSIVGTYSKIMHKSILC 314
Qy      324 FPDDPKVSSDFLLIOSLGCQKERLKEG---LCHPFF--SKIDWNNIRNSPPPVPT 378
Db      315 FPEDAEISKAKULICAFLLDREVRLGNGVEIRQHPFFKNDQWMDNIRETAPVVE 374
Qy      379 LKSDDTNSNDEBPKNW--VSSPCQLSPSGFSEELPFVGFYSKALGILGSESVVG 437
Db      375 LSSDIOSSNDPIDEDDGADVETEPPI--PRAFVGNQLPFGIFYRRENLLLS----- 423
Qy      438 LDPSP--KTSMEKKLLIKSKELQDSQDKCHKMOEOTRLHRRVSEVAVLSQKEVELKA 495
Db      424 -DPSCHENDSTQSR--KNESQELQKLYTBELHS-----EKQA 462
Qy      496 SETQRLSEODLATYITECSLSKSLQAREVSOE-----DDKALQLHDIREQSRKL 549
Db      463 KEE---LEQ-----KCKSVNTRLEKTAKELEBEITLAKSVESALRQL--ERRKAL 507
Qy      550 QEIKQSYQAOVBEWRLMNNQLEEDLVASRRSDLYSELSRESRLAAEFKRAKATBCQHK 609
Db      508 LQHKNAEYQKRAHDAEDKKNLENDVNSLKQULEDLKKKQNSQISTE---KVNLQLRQ 563
Qy      610 LKAKQOGRKEVEGYAKLEKINAEQKIQLOLEKLEKAVAKASTAEATELLONTROAKERA 669
Db      564 LDEFNALLRTESDPTAALRLTQABSSKQIQOULSN-----NRDQDKNCLLETA 612
Qy      670 ERELEK-LQNRDSESEGRKKLVEAEERRSLNKVRLTEMERENRLKDDIQTKSQOI 728
Db      613 KKLIEKRFILQALBESRRDRTHGSEIINDIGRIGLEBDLKNGKILAKVLEKRLQ 672
Qy      729 QQWADKLELEBKREAOVSAQ--HLEVHLKQKE-QHYEEKIVLD-NQIKQDL--ADKE 782
Db      673 QERFTD-LEKESNMEDMTYQVKIQQSLEGEAEHKAATKARLADKKIYESIEEAKSE 731
Qy      783 TLNNMQRHEEAHEKGI---LSEQAMINAMDSKIRSLQRIVEISEANKLAANSLSL 839
Db      732 AKMEMEKKLEEBRTLKQVENILLAEKRGSLDDCDLQSQQKINEL----- 778
Qy      840 TORNMKAQOEMISELRQOKFYLETQAGKLEAQNRLKEBQKLEKISHQSHSDKNRLLEETR 899
Db      779 ----LKQKDVANDVNLTLKIEQETQKCLTQNDLKMQOQVNTLKMSRK-QUKQENNH 833
Qy      900 LREVSLEHEQKLEKRLQLELOLSQRESQLTALQADAALESQURQAKTELEETT-- 957

```

Db 834 LMEKKNLEKONAEIRKRRQADQOMKELOQULABOYFSTLYTKYVRELKEECEEKTKL 893
Qy 958 -AAEAEELIATARDEIQRRFDALRNSCTVITDLEBOLNQTEDBNELNNONYLSKOL 1016
Db 894 GKELOQKQKQLODBRDSIAAQLE-----ITLTQDSEQLANSIAEBEQSDLEKEKIM 945
Qy 1017 DEAGNADIVQLASEVDHILAREITEREMQTSQKQTMALAKTCTMLEBQVMLELND 1076
Db 946 KE-----LEIKEMARHKQELTEKDATIASLETRTLTSDVANIAN--EKEELNN 994
Qy 1077 ELLEKQRQWEMARSVLGDE-----KSQFECRVELQRMIDTEKQSRARADQRITESRQ 1129
Db 995 KLKQVQER---LRLNDEEISAAAIKQFE-----IKYOK 1086
Qy 1130 VVELAVENKKAELIALQALKEQKLKAKESLSDKLNLEKKIAMLEKNAISIQQLTERE 1189
Db 1022 -----KQLTERTLKTQAVNKLAEIMRKSPKRGKNDTVRRKEKENRK 1065
Qy 1190 LKQRLBERQALQOQMDLQKNIHRLTQGLQELDRADLKTESQDLEYQLENTQVLYSH 1249
Db 1066 LHMEIKSRERKLTQOM-----IKYOK 1086
Qy 1250 EKVMEGTISQOTKLIDPLQAKMDQPAKKKVPLOYNELKLAEKAKARCAELEALQKT 1309
Db 1087 ELNEMQOIAEESQI-----RIELQNTLDSKSDITQQLSOLAL 1126
Qy 1310 RIELSARAEEMARKATDPHPSTPATARQOIASAIVRSEPHOPSANSLAPSSRKE 1369
Db 1127 HIGDSS-----SISGSGDAHAD 1145
Qy 1370 SSTEEFSRLKEMHNHPIRFVNGLMNATKCAVCLDTHFGQASKLCEQVGMCHPK 1429
Db 1146 DGFPES----- 1151
Qy 1430 CSTCLPATCGLPAYATHTEAFGRDKMNSPGLQTKEPSSSLHEGMMKVPRNNKGGQG 1489
Db 1152 -----RLGMLSLFVRNNYTKFG 1169
Qy 1490 WDRKYIVLEGSKYLIDNEAREAGQRP 1516
Db 1170 WVKKYIVSSSKILLFYDSEQDKQSNP 1196

RESULT 13
US-09-916-204-4
; Sequence 4, Application US/09916204
; Patent No. 6638745
; GENERAL INFORMATION:
; APPLICANT: WEI, Ming-Hui et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; FILE REFERENCE: CL001164CIP
; CURRENT APPLICATION NUMBER: US/09/916,204
; CURRENT FILING DATE: 2001-07-24
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 251
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-09-916-204-4

Query Match 11.3%; Score 1183; DB 4; Length 251;
Best Local Similarity 90.8%; Pred. No. 4.9e-54;
Matches 228; Conservative 12; Mismatches 11; Indels 0; Gaps 0;

Qy 1 MLKFKYGARNPLDGAAPIASRASRLNLFQGRKPFMTQOQMSPLSRBGIDLALFVLF 60
Db 1 MLKFKYGVNRPSEASAPPIASRASRLNLFQGRKPLMTQOQMSALSRBGIDLALFVLF 60
Qy 61 ECGQPALMKIKHVSNFVRKYSDDTIAELQELQPSAKDFEVRSLVCGGFAEVQVVRKATG 120
Db 61 ECGQPALMKIKHVSNFVRKYSDDTIAELQELQPSAKDFEVRSLVCGGFAEVQVVRKATG 120

Db 61 ECGQPALMKIKHVSNFVRKYSDDTIAELQELQPSVRDFEVRSLVCGGFAEVQVVRKATG 120
Qy 121 DIYAMVKMKKALLAEOVSFFEEERNILSRSTSPMIPOLQYAFODKNHLYLMEYOPGG 180
Db 121 DVIYAMKIMKKAALRAEOVSFFEEERNILSQSTSPMIPOLQYAFODKNHLYLMEYOPGG 180
Qy 181 DLISLNRVEDQDENLQFYLAELILAVSVHLMGVYHRDIPENILVDRGTGHIKLVDF 240
Db 181 DLISLNRVEDQDENLQFYLAELILAVSVHQMVGVRHDIPENILVDRGTGHIKLVDF 240
Qy 241 GSAAKNNSNM 251
Db 241 GSAAKNNSNV 251

RESULT 14
US-10-282-048-4
; Sequence 4, Application US/10282048
; Patent No. 6692948
; GENERAL INFORMATION:
; APPLICANT: WEI, Ming-Hui et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; FILE REFERENCE: CL001164CIP-DIV
; CURRENT APPLICATION NUMBER: US/10/282,048
; CURRENT FILING DATE: 2002-10-29
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 251
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-10-282-048-4

Query Match 11.3%; Score 1183; DB 4; Length 251;
Best Local Similarity 90.8%; Pred. No. 4.9e-54;
Matches 228; Conservative 12; Mismatches 11; Indels 0; Gaps 0;

Qy 1 MLKFKYGARNPLDGAAPIASRASRLNLFQGRKPFMTQOQMSPLSRBGIDLALFVLF 60
Db 1 MLKFKYGVNRPSEASAPPIASRASRLNLFQGRKPLMTQOQMSALSRBGIDLALFVLF 60
Qy 61 ECGQPALMKIKHVSNFVRKYSDDTIAELQELQPSAKDFEVRSLVCGGFAEVQVVRKATG 120
Db 61 ECGQPALMKIKHVSNFVRKYSDDTIAELQELQPSVRDFEVRSLVCGGFAEVQVVRKATG 120
Qy 121 DIYAMVKMKKALLAEOVSFFEEERNILSRSTSPMIPOLQYAFODKNHLYLMEYOPGG 180
Db 121 DVIYAMKIMKKAALRAEOVSFFEEERNILSQSTSPMIPOLQYAFODKNHLYLMEYOPGG 180
Qy 181 DLISLNRVEDQDENLQFYLAELILAVSVHLMGVYHRDIPENILVDRGTGHIKLVDF 240
Db 181 DLISLNRVEDQDENLQFYLAELILAVSVHQMVGVRHDIPENILVDRGTGHIKLVDF 240
Qy 241 GSAAKNNSNM 251
Db 241 GSAAKNNSNV 251

RESULT 15
US-09-916-204-5
; Sequence 5, Application US/09916204
; Patent No. 6638745
; GENERAL INFORMATION:
; APPLICANT: WEI, Ming-Hui et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; FILE REFERENCE: CL001164CIP
; CURRENT APPLICATION NUMBER: US/09/916,204
; CURRENT FILING DATE: 2001-07-24
; NUMBER OF SEQ ID NOS: 6

SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 251
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-916-204-5

Query Match 11.2%; Score 1180; DB 4; Length 251;
Best Local Similarity 90.4%; Fred. No. 7e-54; 9; Indels 0; Gaps 0;
Matches 227; Conservative 15; Mismatches 9;

Qy	1	MLKFEYGARNDPLDGAAPPIASRASRLNLFQGKPPFTQOOMSPISREGIIDLFLVLE	60
Db	1	MLKFEYGVNRPPEASASEPIASRASRLNLFQGKPPFTQOOMSALSRGKLDLFLVLE	60
Qy	61	ECSQPALMKIKHVSNFVRKYSDTIAELQELQPSAKDFEVRSLVGCCHPAEYQVVRKATG	120
Db	61	ECSQPALMKHGVSSFVQKYSDTIAELRELQPSARDFEVRSLVGCCHPAEYQVVRKATG	120
Qy	121	DIYAMKWKKKALLAQEQVSFPEEERNILSRSTSPWIPOLQYAFQDKNHLVLVMEYQPG	180
Db	121	DVYAMKIKKKALLAQEQVSFPEEERNILSRSTSPWIPOLQYAFQDKNHLVLVMEYQPG	180
Qy	181	DLSTLNRYEDQDENLQFYLAELILAVSHVHLMGYVHRDIKPNILVDRGTGHIKLVDF	240
Db	181	DFLSLNRYEDQDENLQFYLAELILAVSHVHLMGYVHRDIKPNILVDRGTGHIKLVDF	240
Qy	241	GSAAKMNSNM 251	
Db	241	GSAAKMNSNKV 251	

Search completed: May 11, 2005, 16:46:48
Job time : 65 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 11, 2005, 16:27:33 ; Search time 227 Seconds
(without alignments)
3499.585 Million cell updates/sec

Title: US-10-791-666-2

Perfect score: 10493
Sequence: 1 MLKFKYGRNPIDAGAAAPR.....AGAVRTPISQVKNWQSSV 2054

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_16Dec04:*
1: geneseqp1980a:*
2: geneseqp1990a:*
3: geneseqp2000a:*
4: geneseqp2001a:*
5: geneseqp2002a:*
6: geneseqp2003a:*
7: geneseqp2003ba:*
8: geneseqp2004a:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	10493	100.0	2054	5	ABR81927 Human kin
2	10486	99.9	2054	5	AAE24150 Human kin
3	10486	99.9	2054	6	AAO26959 Human CRT
4	10464.5	99.7	2055	8	ADJ96610 Human cIt
5	10458.5	99.7	2053	5	ABG78363 RHO/RAC-1
6	10458.5	99.7	2053	6	ADA05642 Human NOV
7	10431.5	99.4	2053	4	AAU03501 Human pro
8	10425.5	99.4	2066	5	ABG78362 Human pro
9	10425.5	99.4	2066	6	ADA05654 Human NOV
10	10422.5	99.3	2053	8	ADN62807 Human NOV
11	10412.5	99.2	2053	8	ADN63228 Human NOV
12	10407.5	99.2	2066	8	ADN62819 Human NOV
13	10109.5	96.3	2055	6	ABP97683 Polypept
14	10109.5	96.3	2055	6	AAO26960 Human CRT
15	10022.5	95.5	2053	5	AAE24079 Human MRP
16	10022.5	95.5	2053	7	ADRF60994 Human CRP
17	10022.5	95.5	2053	8	ADG89100 Human aaso
18	9916	94.5	1958	5	ABR81928 Human turo
19	8161	77.8	1641	7	ADN62804 Mouse cIt
20	7851.5	74.8	1619	7	ADD48584 Rat Prote
21	7851.5	74.8	1619	7	ADD46616 Rat Prote
22	6575	62.7	1286	3	AAH43359 Human ORF
23	6575	62.7	1286	6	AAO26961 Human CRT
24	6575	62.7	1286	7	ADD48586 Human Pro
25	6575	62.7	1286	7	ADD46618 Human Pro

26	5037	48.0	999	4	ABR11117 Human RHO
27	4871	46.4	940	7	ADD89967 Human Can
28	4601	43.8	883	8	ADM72201 Human TAS
29	4501	42.9	910	4	AAU31443 Novel hum
30	4303	41.0	853	7	ADG09957 Novel pro
31	4115	39.2	832	6	ADA05646 Human NOV
32	4115	39.2	832	6	ADN62811 Human NOV
33	4097.5	39.0	847	6	ADA05648 Human NOV
34	4093.5	39.0	847	6	ADN62813 Human NOV
35	3244	30.9	623	6	ADA05650 Human NOV
36	3226.5	30.7	638	6	ADA05652 Human NOV
37	3226.5	30.7	638	8	ADN62817 Human NOV
38	3219	30.7	623	8	ADN62815 Human NOV
39	3165	30.2	623	6	ADA05644 Human NOV
40	3104.5	29.6	620	8	ADN62809 Human NOV
41	2443	23.3	495	6	ABP97681 Amino ac1
42	2443	23.3	497	6	ABP97687 Amino ac1
43	2440	23.3	497	6	ABU10126 Novel hum
44	2440	23.3	497	7	AAE39504 Human kin
45	2440	23.3	497	8	ADA05692 Human kin

ALIGNMENTS

RESULT 1					
ABR81927	ID	ABR81927 standard; protein; 2054 AA.			
XX	AC	ABR81927;			
XX	DT	10-OCT-2002 (first entry)			
XX	DE	Human kinase #1.			
XX	KW	Human; kinase; enzyme; serine-threonine kinase; nocotropic; cytosolic;			
XX	KM	Citron rho-interacting kinase; gene therapy; mental disorder; cancer.			
XX	OS	Homo sapiens.			
XX	PN	WO200259325-A2.			
XX	PD	01-AUG-2002.			
XX	PF	20-DEC-2001; 2001WO-US050497.			
XX	PR	27-DEC-2000; 2000US-0258335P.			
XX	PS	(LEXI-) LEXICON GENETICS INC.			
XX	PT	Yu X, Miranda M, Fiddie CJ;			
XX	DR	WPI; 2002-599796/64.			
XX	PT	N-PSDB; ABQ78870.			
XX	PT	Novel polynucleotide encoding human proteins that are structurally			
XX	PT	similar to animal kinases, useful for drug screening, diagnosis, in gene			
XX	PT	therapy of disorders and diseases e.g. cancer and pharmacogenomic			
XX	PT	applications.			
XX	PS	Claim 2; Page 39-43; 50pp; English.			
XX	CC	The invention relates to a novel human protein that shares structural			
XX	CC	similarity with animal kinases, including serine-threonine kinases,			
XX	CC	particularly Citron rho-interacting kinases. The proteins of the			
XX	CC	invention have nocotropic and cytosolic activity. The polynucleotides may			
XX	CC	have a use in gene therapy. The encoded novel polypeptides are useful for			
XX	CC	generating antibodies, as reagents in diagnostic assays, for identifying			
XX	CC	other cellular gene products related to NHP and as reagents in assays for			
XX	CC	screening for compounds that are useful in the treatment of mental,			
XX	CC	biological or medical disorders and diseases including cancer. The			
XX	CC	sequence represents a novel human kinase of the invention			

Seq	Sequence	2054 AA;
	Query Match	100.0%; Score 10493; DB 5; Length 2054;
	Best Local Similarity	100.0%; Pred. No. 0;
	Matches 2054;	Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	1 MKLFYGAANPLDAGAAEPDASASRLNFFQGGKPPMTQOQMSPLREGIILDALFVLE	60
DB	1 MKLFYGAANPLDAGAAEPDASASRLNFFQGGKPPMTQOQMSPLREGIILDALFVLE	60
QY	61 ECSQALMKIKVSNFVRKYSDTIAELQELQPSAKDFEVRSLVGGCFHAEVQVEEKATG	120
DB	61 ECSQALMKIKVSNFVRKYSDTIAELQELQPSAKDFEVRSLVGGCFHAEVQVEEKATG	120
QY	121 DIYAMKVMKKALLAQEOVSPFEEERNILSRSTSPMIPOLQYAFODKNHLYLMEXQPG	180
DB	121 DIYAMKVMKKALLAQEOVSPFEEERNILSRSTSPMIPOLQYAFODKNHLYLMEXQPG	180
QY	181 DLLSLINRYEDQDENLQFYLAELLIAVSHVMGYVRDIPKENIIVDRGHIKLVDF	240
DB	181 DLLSLINRYEDQDENLQFYLAELLIAVSHVMGYVRDIPKENIIVDRGHIKLVDF	240
QY	241 GSAAKNNSKVMYNAKLPIGTPTYMAPEVLTVMNGDGKGTGGLDCDWMVGVIAEMITYGR	300
DB	241 GSAAKNNSKVMYNAKLPIGTPTYMAPEVLTVMNGDGKGTGGLDCDWMVGVIAEMITYGR	300
QY	301 SPFAGTSARTFNINNFQRFKFPDDPVYSDFDLILQSLCGQKERLKFEGLCCHPFF	360
DB	301 SPFAGTSARTFNINNFQRFKFPDDPVYSDFDLILQSLCGQKERLKFEGLCCHPFF	360
QY	361 SKIDMNNIRNSPPPYPTIKSDDDTNSPDEPKNSVSSPCQLSPSGSGSEELPVGVS	420
DB	361 SKIDMNNIRNSPPPYPTIKSDDDTNSPDEPKNSVSSPCQLSPSGSGSEELPVGVS	420
QY	421 YSKALGIIGRSVYSGLDSPAKTSMKKLTKSKELQSDQCKHMEQEMTRILHRYS	480
DB	421 YSKALGIIGRSVYSGLDSPAKTSMKKLTKSKELQSDQCKHMEQEMTRILHRYS	480
QY	481 EYEAVALSKEVELKASETQSRSLLEODLATYITECSSLKSLBOARMEVSQEDDKALQLLH	540
DB	481 EYEAVALSKEVELKASETQSRSLLEODLATYITECSSLKSLBOARMEVSQEDDKALQLLH	540
QY	541 DIRBOSRKLQETKEQHYQAQVEEMRLMNOLEBDLVASARRSDLYSELSRESRLAAEEK	600
DB	541 DIRBOSRKLQETKEQHYQAQVEEMRLMNOLEBDLVASARRSDLYSELSRESRLAAEEK	600
QY	541 DIRBOSRKLQETKEQHYQAQVEEMRLMNOLEBDLVASARRSDLYSELSRESRLAAEEK	600
DB	541 DIRBOSRKLQETKEQHYQAQVEEMRLMNOLEBDLVASARRSDLYSELSRESRLAAEEK	600
QY	601 RKATSCOHKLAKXQKQPEVGEYAKLEKINAQOQKIQELOBKLEKAYKASTEATELHQ	660
DB	601 RKATSCOHKLAKXQKQPEVGEYAKLEKINAQOQKIQELOBKLEKAYKASTEATELHQ	660
QY	661 NIROAKERAERLEKLQNRDSESGIRKKLVEAEERRHSLBNKVKLLETMERRENRKLD	720
DB	661 NIROAKERAERLEKLQNRDSESGIRKKLVEAEERRHSLBNKVKLLETMERRENRKLD	720
QY	721 IOTKSQOQIQWADKILIEEKEHREAQVSAQHLVHLKQKEQHYBEKIKYLDNQIKKDLAD	780
DB	721 IOTKSQOQIQWADKILIEEKEHREAQVSAQHLVHLKQKEQHYBEKIKYLDNQIKKDLAD	780
QY	781 KETLNNMORHEEAEHEKGIISQOKAMTNAMDSKIRSLRORIVELSEANKLAANSSLT	840
DB	781 KETLNNMORHEEAEHEKGIISQOKAMTNAMDSKIRSLRORIVELSEANKLAANSSLT	840
QY	841 QRMNKAQOEEMISELNOQKFYLETOAGKLEAONRKLEBOLEKISHODSHDKNLLLEETL	900
DB	841 QRMNKAQOEEMISELNOQKFYLETOAGKLEAONRKLEBOLEKISHODSHDKNLLLEETL	900
QY	901 REVSLEHEBQKLELRQLTQLSLQERESQUTALQAAAPALESQLRQKTELEETTABA	960
DB	901 REVSLEHEBQKLELRQLTQLSLQERESQUTALQAAAPALESQLRQKTELEETTABA	960
QY	961 EEELQALTARDEIQRKFPALRNSCTVITDLEBQNLQETEDNAELNNQFYLSKQLDERS	1020
DB	961 EEELQALTARDEIQRKFPALRNSCTVITDLEBQNLQETEDNAELNNQFYLSKQLDERS	1020

QY	1021 GANDEIVQLRSEVYDLRREITREMOVLTSQOKOTMEALKTCTMLEBOVMDLEALDELLE	1080
DB	1021 GANDEIVQLRSEVYDLRREITREMOVLTSQOKOTMEALKTCTMLEBOVMDLEALDELLE	1080
QY	1081 KERQWEAMRSVLDGEKSQPECEVRELQRMILDEKOSRARADQRTIESQVVELAVKEHKA	1140
DB	1081 KERQWEAMRSVLDGEKSQPECEVRELQRMILDEKOSRARADQRTIESQVVELAVKEHKA	1140
QY	1141 EILALQALKEQKLAESISDKLNDLEKKHMLBNANASLOQKLETERELKORLLEBOAK	1200
DB	1141 EILALQALKEQKLAESISDKLNDLEKKHMLBNANASLOQKLETERELKORLLEBOAK	1200
QY	1201 LQOQMDLQKNHIFRLTQGLQELADLADILKTRSLLEYQLENIQVLYSHKQKMGSTISQ	1260
DB	1201 LQOQMDLQKNHIFRLTQGLQELADLADILKTRSLLEYQLENIQVLYSHKQKMGSTISQ	1260
QY	1261 QTKLIDFLQAKMDQPAKKKKVPLQYNELKLALEKKAACAELEALOKTRIELRSAREBA	1320
DB	1261 QTKLIDFLQAKMDQPAKKKKVPLQYNELKLALEKKAACAELEALOKTRIELRSAREBA	1320
QY	1321 AHRKATDHPHSTPATAPAOQIAMSALVRSPEHQPSAMSLAPSSRRKESSTPEEFRRLL	1380
DB	1321 AHRKATDHPHSTPATAPAOQIAMSALVRSPEHQPSAMSLAPSSRRKESSTPEEFRRLL	1380
QY	1381 KERMEHNIIPHRPNVGLNRPATYKCAVCLDTTFHFGQASKLEQOVNCHKXCSTCLPATGL	1440
DB	1381 KERMEHNIIPHRPNVGLNRPATYKCAVCLDTTFHFGQASKLEQOVNCHKXCSTCLPATGL	1440
QY	1441 PAEYATHFTFPAACRQKMSPGLOTKEPSSSLHLEGMMKVPRRNKKGOQMDRKYIVLBS	1500
DB	1441 PAEYATHFTFPAACRQKMSPGLOTKEPSSSLHLEGMMKVPRRNKKGOQMDRKYIVLBS	1500
QY	1501 KVLIDNEAREAGQRPVEEFBELCLPDGVSJHGAVASLANTADADVPYILKQMSHPRT	1560
DB	1501 KVLIDNEAREAGQRPVEEFBELCLPDGVSJHGAVASLANTADADVPYILKQMSHPRT	1560
QY	1561 TCWPGRTIYLLAPSPDYQQRWYTTALESVYAGRVREKAEADAKILGNSILTKBEDDRD	1620
DB	1561 TCWPGRTIYLLAPSPDYQQRWYTTALESVYAGRVREKAEADAKILGNSILTKBEDDRD	1620
QY	1621 MNCTLPFSQOVVLYGTEBGLVALANVLYKNSLTVPGVIGAVFOIYITKDEKLMIAGEERA	1680
DB	1621 MNCTLPFSQOVVLYGTEBGLVALANVLYKNSLTVPGVIGAVFOIYITKDEKLMIAGEERA	1680
QY	1681 LCLVDVKVKQKSLAQSHLPAPDISPNIPEAVKGCHLFGAGKIENGGLCIQAMPKVVYL	1740
DB	1681 LCLVDVKVKQKSLAQSHLPAPDISPNIPEAVKGCHLFGAGKIENGGLCIQAMPKVVYL	1740
QY	1741 RYNNELSKYCIKKEIETSEBPCSHFTNYSLIGNNKYEIDMKQYTLBEFLDKNDHSLA	1800
DB	1741 RYNNELSKYCIKKEIETSEBPCSHFTNYSLIGNNKYEIDMKQYTLBEFLDKNDHSLA	1800
QY	1801 PAVFAASSNPFSVIVQVNSAGQREYLLCHFEFVFPDVSIGRRRTDLMKSRILPLAFA	1860
DB	1801 PAVFAASSNPFSVIVQVNSAGQREYLLCHFEFVFPDVSIGRRRTDLMKSRILPLAFA	1860
QY	1861 YREPLYTFHENSLEVIETIOARSSAGTAPARAYLDIPNRYILGPALISSGAILYASSYQKL	1920
DB	1861 YREPLYTFHENSLEVIETIOARSSAGTAPARAYLDIPNRYILGPALISSGAILYASSYQKL	1920
QY	1921 RYICCKGNLVESGTEHHRGPSTSRSSPNKKGAPPTYNEHITKRVASSAPAPGPGSHPREP	1980
DB	1921 RYICCKGNLVESGTEHHRGPSTSRSSPNKKGAPPTYNEHITKRVASSAPAPGPGSHPREP	1980
QY	1981 STPHRYRGRGTRILRDQSPGAPRLERKSPGRILSTRERRSPARLFEDESSRGLPAGAVRT	2040
DB	1981 STPHRYRGRGTRILRDQSPGAPRLERKSPGRILSTRERRSPARLFEDESSRGLPAGAVRT	2040
QY	2041 PLSQVNAKWDOSV 2054	
DB	2041 PLSQVNAKWDOSV 2054	

```
RESULT 2
AAE24150
ID AAE24150 standard; protein; 2054 AA.
XX
AC AAE24150;
XX
DT 23-SEP-2002 (first entry)
XX
DE Human kinase (PKIN)-21 protein.
XX
KW Human; kinase; PKIN; cancer; immune system disorder; atherosclerosis;
KW acquired immune deficiency syndrome; AIDS; Addison's disease; allergy;
KW asthma; multiple sclerosis; psoriasis; arteriosclerosis; cirrhosis;
KW development; hepatitis; cardiovascular; hypertension; drug screening;
KW myocardial infarction; Goodpasture's syndrome; lipid disorder; growth;
KW fatty liver; Gaucher's disease; Niemann-Pick's disease; anorectic;
KW hypercholesterolaemia; obesity; gene therapy; cytostatic; anti-HIV;
KW neuroprotective; hepatotropic; hypotensive; cardiant; nephrotropic;
KW hyperlipidaemia; enzyme.
XX
OS Homo sapiens.
XX
FH Key
FH Domain
FT /note="Location/Qualifiers
FT /note="97..360
FT /note="Eukaryotic protein kinase domain"
FT /note="98..241
FT /note="Protein kinase domain"
FT /note="99..349
FT /note="Protein kinase domain"
FT /note="101..241
FT /note="Protein kinase domain"
FT /note="102..241
FT /note="Protein kinase domain"
FT /note="249..349
FT /note="Protein kinase domain"
FT /note="258..445
FT /note="Protein kinase domain"
FT /note="258..349
FT /note="Protein kinase domain"
FT /note="534..542
FT /note="Domain found in NIK1-like kinase"
FT /note="854..875
FT /note="Leucine zipper pattern"
FT /note="891..933
FT /note="Domain found in NIK1-like kinase"
FT /note="964..975
FT /note="Domain found in NIK1-like kinase"
FT /note="991..1012
FT /note="Leucine zipper pattern"
FT /note="1015..1067
FT /note="Domain found in NIK1-like kinase"
FT /note="1057..1078
FT /note="Leucine zipper pattern"
FT /note="1159..1180
FT /note="Leucine zipper pattern"
FT /note="1217..1255
FT /note="Domain found in NIK1-like kinase"
FT /note="1388..1434
FT /note="Domain found in NIK1-like kinase"
FT /note="1390..1438
FT /note="Phorbol esters/diacylglycerol binding site"
FT /note="1403..1466
FT /note="Phorbol esters/diacylglycerol binding site"
FT /note="1471..1590
FT /note="PH domain"
FT /note="1619..1916
FT /note="CNH domain"
FT /note="1759..1802
FT /note="Domain found in NIK1-like kinase"
FT /note="1819..1831
FT /note="Domain found in NIK1-like kinase"
FT /note="1851..1880
FT Domain
```

```
FT /note="Domain found in NIK1-like kinase"
XX
XX WO200233099-A2.
XX
XX 25-APR-2002.
XX
XX 20-OCT-2001; 2001WO-US047728.
XX
XX 20-OCT-2000; 2000US-0242410P.
XX
XX 27-OCT-2000; 2000US-0244068P.
XX
XX 09-NOV-2000; 2000US-0245708P.
XX
XX 03-NOV-2000; 2000US-0247672P.
XX
XX 16-NOV-2000; 2000US-0249565P.
XX
XX 22-NOV-2000; 2000US-0252730P.
XX
XX 01-DEC-2000; 2000US-0250807P.
XX
XX (INCY-) INCYTE GENOMICS INC.
XX
XX Gururajan R, Baughn MR, Walla NK, Elliott VS, Xu Y, Arvizu C,
XX Yao MG, Ramkumar J, Ding L, Tang YF, Hatella AD, Nguyen DB,
XX Gandhi AR, Lu Y, Yue H, Burford N, Bandman O, Tribouley CM, Lal PG,
XX Recipon SA, Lu DM, Borowsky ML, Thornton M, Swarnaker A,
XX Thangavelu K, Khan FA, Ison CH;
XX WPI; 2002-454603/48.
XX
XX N-PSDB; AAD38864.
XX
XX New human kinase polypeptide, for diagnosing, preventing and treating
XX cancer, immune system disorder, growth and development disorders,
XX cardiovascular disorders and lipid disorders.
XX
XX Claim 1; Page 177-182; 210pp; English.
XX
XX The invention relates human kinases (PKIN) and their corresponding
XX nucleic acid sequences. PKIN and its DNA are useful for diagnosing,
XX treating and preventing cancer, an immune system disorder (e.g., acquired
XX immune deficiency syndrome (AIDS), Addison's disease, allergy, asthma,
XX atherosclerosis, multiple sclerosis, psoriasis), disorders affecting
XX growth and development (e.g., arteriosclerosis, cirrhosis, hepatitis),
XX cardiovascular disorder (e.g., hypertension, myocardial infarction,
XX Goodpasture's syndrome), and a lipid disorder (e.g., fatty liver,
XX Gaucher's disease, Niemann-Pick's disease, hypercholesterolaemia,
XX hyperlipidaemia, obesity), and for assessing the effects of exogenous
XX compounds. Anti-PKIN antibody is useful in a diagnostic test for a
XX condition or a disease associated with the expression of PKIN in a
XX biological sample. A composition comprising PKIN or an agonist or
XX antagonist of PKIN is useful for treating a disease or condition
XX associated with decreased or increased expression of functional PKIN.
XX PKIN is useful in a number of drug screening techniques and to analyse
XX the proteome of a tissue or cell type. PKIN DNA is useful for creating
XX knockin humanised animals or transgenic animals to model human diseases,
XX and in somatic or germ-line gene therapy. The present sequence is human
XX PKIN protein
XX
XX Sequence 2054 AA:
XX
XX Query Match 99.9%; Score 10486; DB 5; Length 2054;
XX Best Local Similarity 99.9%; Pred. No. 0;
XX Matches 2052; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
XX
XX 1 MLKFKYGAAMPDAGAEPIASRASRLNLFPGKPPFTQOQSPLSRGILDALFVLE 60
XX |
XX 1 MLKFKYGAAMPDAGAEPIASRASRLNLFPGKPPFTQOQSPLSRGILDALFVLE 60
XX |
XX 61 ECSQPALMKIKIVSNFVKYSDPTIAELQPSADFEVRSIVGCGHFAEVGVREKATG 120
XX |
XX 61 ECSQPALMKIKIVSNFVKYSDPTIAELQPSADFEVRSIVGCGHFAEVGVREKATG 120
XX |
XX 121 DIYAMKMKKALLAOBOVSFEERNTLSRSTSPWIOLOVAFODKXNLYLWMEYOPCG 180
XX |
XX 121 DIYAMKMKKALLAOBOVSFEERNTLSRSTSPWIOLOVAFODKXNLYLWMEYOPCG 180
XX |
XX 181 DLSLNLRYEDQDENLIGFYLAELILAVSVHLMGVYVARDIKENILVDRGTIKLVDF 240
XX |
```

```

Db 181 DILSLNLYEDLDNLIOFYLAELILAVSHVLMGYHARDIKPENILVDRIGHIKLVDL 240
Qy 241 GSAAKNSNKNWAKNLP1GTPDYMAPEVLTVANGDGKGTGYGLDCDMSVGIAYEM1YGR 300
Db 241 GSAAKNSNKNWAKNLP1GTPDYMAPEVLTVANGDGKGTGYGLDCDMSVGIAYEM1YGR 300
Qy 301 SPFAAGTSARPTNNIMNFQRF1KPPDDPKVSSDFDL1D1SLCGQKRLKFGGLCCHPF 360
Db 301 SPFAAGTSARPTNNIMNFQRF1KPPDDPKVSSDFDL1D1SLCGQKRLKFGGLCCHPF 360
Qy 361 SKIDNNN1RNSPPVPPTLKSDDDTSNPEPEKNSWSSPCQ1SPSGSGEELPFVGS 420
Db 361 SKIDNNN1RNSPPVPPTLKSDDDTSNPEPEKNSWSSPCQ1SPSGSGEELPFVGS 420
Qy 421 YSKALGILGRSESVSGLDSPAKTSSMEKLL1KSKELQDSQDKCHKMEQENTRLHRRVS 480
Db 421 YSKALGILGRSESVSGLDSPAKTSSMEKLL1KSKELQDSQDKCHKMEQENTRLHRRVS 480
Qy 481 EYEAVALSQEVELKASETORSLLEODLATYITEGSLKRSLEQAMENVSQEDDKALQILH 540
Db 481 EYEAVALSQEVELKASETORSLLEODLATYITEGSLKRSLEQAMENVSQEDDKALQILH 540
Qy 541 DIRESRKLQETKEQYQAVEEMRLMMQLEED1VSARRSDLYESELRESRLAABEFK 600
Db 541 DIRESRKLQETKEQYQAVEEMRLMMQLEED1VSARRSDLYESELRESRLAABEFK 600
Qy 601 RKATCQHLKAKQKQKPEVEYAKLEKINAQOLK1QEOLEKLEKAVKASTEATELLO 660
Db 601 RKATCQHLKAKQKQKPEVEYAKLEKINAQOLK1QEOLEKLEKAVKASTEATELLO 660
Qy 661 N1RQAKERARELEK1QNRDESSSEGI1RKKLVABEERRSHLENKVRLETMERERELKXD 720
Db 661 N1RQAKERARELEK1QNRDESSSEGI1RKKLVABEERRSHLENKVRLETMERERELKXD 720
Qy 721 IOTKSQOI1QWADK1LELEKHEKREAVSQH1EYHLKQKEQHYEEKIKYLDNQIKKDLAD 780
Db 721 IOTKSQOI1QWADK1LELEKHEKREAVSQH1EYHLKQKEQHYEEKIKYLDNQIKKDLAD 780
Qy 781 KETLENMMQRRHEEBAHEKGI1SEOKAM1NADSKIR1SEQR1VELSEANKLAANS1FT 840
Db 781 KETLENMMQRRHEEBAHEKGI1SEOKAM1NADSKIR1SEQR1VELSEANKLAANS1FT 840
Qy 841 QRMNKAQSEMI1SEL1FOQKPYLETQAGKLEAQRKLEBQLEK1SHODSHDNKFLLEETRL 900
Db 841 QRMNKAQSEMI1SEL1FOQKPYLETQAGKLEAQRKLEBQLEK1SHODSHDNKFLLEETRL 900
Qy 901 REVSLHEBQKLEKRLQTELO1S1OERESQ1TAL1QAAAL1ESQ1RQAKTELEETTA 960
Db 901 REVSLHEBQKLEKRLQTELO1S1OERESQ1TAL1QAAAL1ESQ1RQAKTELEETTA 960
Qy 961 EEE1QALTAHRDE1QRKFDALNSCTV1TDL1EEO1NQLTEDNA1ELNNQNFYLSKQ1DEAS 1020
Db 961 EEE1QALTAHRDE1QRKFDALNSCTV1TDL1EEO1NQLTEDNA1ELNNQNFYLSKQ1DEAS 1020
Qy 1021 GANDE1VOL1RSEVDH1LRREITEREMQ1TSQOKTMEAL1KTTCTM1LBEOVMD1EALNDEL1E 1080
Db 1021 GANDE1VOL1RSEVDH1LRREITEREMQ1TSQOKTMEAL1KTTCTM1LBEOVMD1EALNDEL1E 1080
Qy 1081 KERQWEARNSV1GDEK1SQPEKCVREL1QRMLDTEK1SRADADR1TESRQV1LAVKENHA 1140
Db 1081 KERQWEARNSV1GDEK1SQPEKCVREL1QRMLDTEK1SRADADR1TESRQV1LAVKENHA 1140
Qy 1141 E1LAL1QOAL1KEOK1LAES1SDL1NDLEK1GAM1EMNARS1LOOK1ETRE1KRL1EBOAK 1200
Db 1141 E1LAL1QOAL1KEOK1LAES1SDL1NDLEK1GAM1EMNARS1LOOK1ETRE1KRL1EBOAK 1200
Qy 1201 LQQQMD1QKNH1FRL1TQGL1OAL1BRAD1LK1TERSD1EYQ1EN1OYV1SH1EKKYMG1ST1SQ 1260
Db 1201 LQQQMD1QKNH1FRL1TQGL1OAL1BRAD1LK1TERSD1EYQ1EN1OYV1SH1EKKYMG1ST1SQ 1260
Qy 1261 QTKL1D1F1QAKMDQPAKKKKVPLQ1YNELK1ALEK1KARCA1E1EALQ1TR1EL1SARBEA 1320

```

```

Db 1261 QTKL1D1F1QAKMDQPAKKKKVPLQ1YNELK1ALEK1KARCA1E1EALQ1TR1EL1SARBEA 1320
Qy 1321 AHRKATDHPHSTPAT1A1QO1AMS1YAS1PEH1QPSAM1LAP1SSRRK1ESS1PEEF1SRL 1380
Db 1321 AHRKATDHPHSTPAT1A1QO1AMS1YAS1PEH1QPSAM1LAP1SSRRK1ESS1PEEF1SRL 1380
Qy 1381 KERMHNN1P1RPNVGLN1AR1KCAV1CLPT1V1PGR1AS1K1E1CQ1V1CH1K1CST1CL1PAT1CGL 1440
Db 1381 KERMHNN1P1RPNVGLN1AR1KCAV1CLPT1V1PGR1AS1K1E1CQ1V1CH1K1CST1CL1PAT1CGL 1440
Qy 1441 PAEYATHTTEA1FCRD1KXNS1PGL1QTK1EPSS1LH1EGMM1KV1PRNNK1GQO1GMD1RKY1VL1EGS 1500
Db 1441 PAEYATHTTEA1FCRD1KXNS1PGL1QTK1EPSS1LH1EGMM1KV1PRNNK1GQO1GMD1RKY1VL1EGS 1500
Qy 1501 KYL1YDN1ARE1AGQ1PVE1FEL1CLPD1GV1SH1GAV1GAS1ELANT1AAD1V1Y1LK1MESH1PT 1560
Db 1501 KYL1YDN1ARE1AGQ1PVE1FEL1CLPD1GV1SH1GAV1GAS1ELANT1AAD1V1Y1LK1MESH1PT 1560
Qy 1561 TCWPGRT1YL1AP1SP1QQRW1TAL1ESV1VAG1R1VRE1KA1EAD1AKL1GNS1LK1E1GDD1RLD 1620
Db 1561 TCWPGRT1YL1AP1SP1QQRW1TAL1ESV1VAG1R1VRE1KA1EAD1AKL1GNS1LK1E1GDD1RLD 1620
Qy 1621 MNCT1PF1SDQ1VVL1GTE1EGL1YAL1NVL1KNS1L1HP1G1GAV1FQ1Y11K1D1EKL1M1AGE1BA 1680
Db 1621 MNCT1PF1SDQ1VVL1GTE1EGL1YAL1NVL1KNS1L1HP1G1GAV1FQ1Y11K1D1EKL1M1AGE1BA 1680
Qy 1681 L1CLV1DK1VK1OS1LAQ1SH1P1AOP1D1SPN1FEAV1KG1CH1L1FGAG1K1ENG1LC11CAAMP1SK1VV1L 1740
Db 1681 L1CLV1DK1VK1OS1LAQ1SH1P1AOP1D1SPN1FEAV1KG1CH1L1FGAG1K1ENG1LC11CAAMP1SK1VV1L 1740
Qy 1741 RYNEML1SKY1C1R1KE1ET1E1SP1C1H1FT1NYS1L1IGN1K1EY1E1DM1KQ1YTL1E1EFL1DK1H1SLA 1800
Db 1741 RYNEML1SKY1C1R1KE1ET1E1SP1C1H1FT1NYS1L1IGN1K1EY1E1DM1KQ1YTL1E1EFL1DK1H1SLA 1800
Qy 1801 PAVFAAS1NSP1SV1VQNS1AGQ1RE1YL1C1HE1F1GV1FYDS1YGR1SR1TDL1KMS1RL1PLA1FA 1860
Db 1801 PAVFAAS1NSP1SV1VQNS1AGQ1RE1YL1C1HE1F1GV1FYDS1YGR1SR1TDL1KMS1RL1PLA1FA 1860
Qy 1861 YREPYL1FT1HNS1LE1V1E1Q1ARSS1AG1PAR1AY1D1PN1RY1GPA1ISG1AY1ASS1YOD1KL 1920
Db 1861 YREPYL1FT1HNS1LE1V1E1Q1ARSS1AG1PAR1AY1D1PN1RY1GPA1ISG1AY1ASS1YOD1KL 1920
Qy 1921 RY1CCK1NVL1VES1GTE1H1RG1ST1SS1PNK1G1PR1YNN1HT1TR1V1SS1PA1R1G1SH1PRE1P 1980
Db 1921 RY1CCK1NVL1VES1GTE1H1RG1ST1SS1PNK1G1PR1YNN1HT1TR1V1SS1PA1R1G1SH1PRE1P 1980
Qy 1981 STPHRYR1GRT1EL1ARD1K1SG1R1PL1ERE1K1SG1IL1STR1RR1SP1AR1L1F1ED1S1R1GL1P1AG1AV1RT 2040
Db 1981 STPHRYR1GRT1EL1ARD1K1SG1R1PL1ERE1K1SG1IL1STR1RR1SP1AR1L1F1ED1S1R1GL1P1AG1AV1RT 2040
Qy 2041 PLSQV1NKV1MDQ1SSV 2054
Db 2041 PLSQV1NKV1MDQ1SSV 2054

RESULT 3
AAO26959
ID AAO26959 standard; protein; 2054 AA.
XX
AC AAO26959;
XX
DT 01-MAY-2003 (first entry)
XX
DE Human CR1K protein sequence, SEQ ID No 2.
XX
KW Anorectic; hypotensive; cardiant; antilipemic; cerebroprotective;
KW antiagout; osteopathic; antiarthritic; cyostatic; antidepressant;
KW immunomodulator; antimanic; tranquiliser; antiparkinsonian; nootropic;
KW neuroprotective; antiinflammatory; antidiabetic; analgesic;
KW human citron rho;rac-interacting kinase; enzyme; CR1K; ameliorating;
KW obesity; comorbidity; cancer; anorexia; cachexia; bulimia;
KW central nervous system disorder; chronic obstructive pulmonary disease;
KW diabetes; pain.

```

XX Homo sapiens.
OS
XX MO2003004523-A1.
PN
XX 16-JAN-2003.
PD
XX 28-JUN-2002; 2002MO-EP007156.
PF
XX 02-JUL-2001; 2001US-0301841P.
PR 11-DEC-2001; 2001US-0338651P.
PR 25-APR-2002; 2002US-0375014P.
XX (FARB) BAYER AG.
PA
XX Zhu Z;
PI
XX MPI; 2003-221576/21.
DR N-PSDB; MAL55214.
DR
XX New human citron rho/rac-interacting kinase (CRIK) polypeptide and
PT polynucleotide, useful in preventing, ameliorating or treating diseases
PT associated with human CRIK dysfunction, e.g. obesity, diabetes or
PT Alzheimer's disease.
XX
XX Claim 1; Fig 2; 237pp; English.

XX The invention relates to an isolated polynucleotide encoding a human
XX citron rho/rac-interacting kinase polypeptide. The isolated
XX polynucleotide comprises a 6165 or 8603 base pair sequence, given in the
XX specification. The human citron rho/rac-interacting kinase (CRIK)
XX polypeptide and polynucleotide are useful in preventing, ameliorating, or
XX treating diseases associated with human CRIK dysfunction such as obesity
XX and obesity-associated comorbidities (e.g. hypertension, coronary artery
XX disease, hyperlipidemia, stroke, gout, osteoarthritis, some types of
XX cancer, including endometrial, breast, prostate and colon cancer),
XX anorexia, cachexia, bulimia, central nervous system disorders (e.g. mood
XX disorders, anxiety disorders, Parkinson's disease or Alzheimer's
XX disease), chronic obstructive pulmonary disease, or diabetes. These can
XX also be used to treat pain associated with the disorders. The human CRIK
XX polypeptide is also useful in diagnostic assays or in genetic testing.
XX The expression vector or the reagent is useful in preparing a medicament
XX for modulating the activity of a human CRIK in a disease, e.g. obesity, a
XX central nervous system disorder, or chronic obstructive pulmonary
XX disease. The fusion protein is useful for generating antibodies against a
XX CRIK polypeptide and for use in various assay systems. The methods are
XX useful in producing and detecting the polynucleotide and polypeptide and
XX in screening for agents that modulate the activity of the human CRIK
XX polypeptide. This sequence represents the human CRIK protein of the
XX invention
XX
XX Sequence 2054 AA;

Query Match 99.9%; Score 10486; DB 6; Length 2054;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2052; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MLKPYGARNPLDGAAPISASRLNLFQKPPPTQOQMSLSREGILDLFVLF 60
DB 1 MLKPYGARNPLDGAAPISASRLNLFQKPPPTQOQMSLSREGILDLFVLF 60
QY 61 ECSOPALMKIKHVSNFVKYSYSDTIAELOPSADPFVRSIVGCGHAEQVYVREKKTG 120
DB 61 ECSOPALMKIKHVSNFVKYSYSDTIAELOPSADPFVRSIVGCGHAEQVYVREKKTG 120
QY 121 DIYAMKVKKKALLAQEOVSFFEEERNILSRSTSPWIDQYAFQDKNHLVLMVEYQCGG 180
DB 121 DIYAMKVKKKALLAQEOVSFFEEERNILSRSTSPWIDQYAFQDKNHLVLMVEYQCGG 180
QY 121 DIYAMKVKKKALLAQEOVSFFEEERNILSRSTSPWIDQYAFQDKNHLVLMVEYQCGG 180
DB 121 DIYAMKVKKKALLAQEOVSFFEEERNILSRSTSPWIDQYAFQDKNHLVLMVEYQCGG 180
QY 181 DLSLNLNRYEDQDLNLIQFYLAELILAVHSVHLMGYVHRDIPKNNIIVDRGTGKLVDF 240
DB 181 DLSLNLNRYEDQDLNLIQFYLAELILAVHSVHLMGYVHRDIPKNNIIVDRGTGKLVDF 240

QY 241 GSAAMNSNKNVNAKLPIGTDPYNAPEVLTVMNGGKGTGYGLDCMWSVGVATYEMITYCR 300
DB 241 GSAAMNSNKNVNAKLPIGTDPYNAPEVLTVMNGGKGTGYGLDCMWSVGVATYEMITYCR 300
QY 301 SPFAGTARTFNNTMNFORFLKPPDDPKVSSDFDLIQLSLCGKERIKFSGLCCHPFF 360
DB 301 SPFAGTARTFNNTMNFORFLKPPDDPKVSSDFDLIQLSLCGKERIKFSGLCCHPFF 360
QY 361 SKIDMNNIRNSPPPEVPTLKSDDDTSNFDPEKNSWSSPCQSPSGSGEELPFVGRS 420
DB 361 SKIDMNNIRNSPPPEVPTLKSDDDTSNFDPEKNSWSSPCQSPSGSGEELPFVGRS 420
QY 421 YSKALGIIGRSBSVSGIDSPAKTSSMEKLLIKSKELQDSQDKCHKMEQENTRLHRRVS 480
DB 421 YSKALGIIGRSBSVSGIDSPAKTSSMEKLLIKSKELQDSQDKCHKMEQENTRLHRRVS 480
QY 481 EYEAVLSQKEVELKASETORSLLEODLATYTTTECSLRSLRQAMWVSQEDKALQLH 540
DB 481 EYEAVLSQKEVELKASETORSLLEODLATYTTTECSLRSLRQAMWVSQEDKALQLH 540
QY 541 DIREQSRKLOEIKEOEYOAYEEMRLMNOLEEDLVASRRSDLYESELRSESLAAEPRK 600
DB 541 DIREQSRKLOEIKEOEYOAYEEMRLMNOLEEDLVASRRSDLYESELRSESLAAEPRK 600
QY 601 RKATECOHKLKAKDQKPEVGEYAKLEKINABQQLKIQLQEKLEKAVKASTATELLQ 660
DB 601 RKATECOHKLKAKDQKPEVGEYAKLEKINABQQLKIQLQEKLEKAVKASTATELLQ 660
QY 661 NITQAKAEAELEGLQNRREDSSEGRKLYAEERRRSLBNKVRLTMEERERRLDD 720
DB 661 NITQAKAEAELEGLQNRREDSSEGRKLYAEERRRSLBNKVRLTMEERERRLDD 720
QY 721 IOTSQOIQOQADKILKEEKGRAOVSQHLVEYLKQEOHYEERIKVLDNQIKKDLAD 780
DB 721 IOTSQOIQOQADKILKEEKGRAOVSQHLVEYLKQEOHYEERIKVLDNQIKKDLAD 780
QY 781 KETLENNMQRBEBAHEKGLISBQKAMINAMDSKIRSLBQRIVELSEANKLAANSUFT 840
DB 781 KETLENNMQRBEBAHEKGLISBQKAMINAMDSKIRSLBQRIVELSEANKLAANSUFT 840
QY 841 QRNKAQOEMISELRQKFFYLETQAKLEPAQNRKLEBOLKESHOHSDKRLLELFTL 900
DB 841 QRNKAQOEMISELRQKFFYLETQAKLEPAQNRKLEBOLKESHOHSDKRLLELFTL 900
QY 901 REVSLEHEQKLEKQOLTELQLSQREBSQTLQAARALBESQLRQKTELESTTAA 960
DB 901 REVSLEHEQKLEKQOLTELQLSQREBSQTLQAARALBESQLRQKTELESTTAA 960
QY 961 EEBIOLATNHRDEIQRKFDALRNSCTVITDLBEOQLNQTEDNABEINMNOFYLSDQDEAS 1020
DB 961 EEBIOLATNHRDEIQRKFDALRNSCTVITDLBEOQLNQTEDNABEINMNOFYLSDQDEAS 1020
QY 1021 GANDEIVOLRSEVDHLRREITEREMOLTSQKQTEALKTTCMLBEOQVMDLEALNDELIE 1080
DB 1021 GANDEIVOLRSEVDHLRREITEREMOLTSQKQTEALKTTCMLBEOQVMDLEALNDELIE 1080
QY 1081 KERQMEARSVLGDPEKQFECGRVRELQMLTEKQSRARQORTESQVVELAYKEKA 1140
DB 1081 KERQMEARSVLGDPEKQFECGRVRELQMLTEKQSRARQORTESQVVELAYKEKA 1140
QY 1141 EILALQALKEQKAKAESISDKLNDLEKGHAMLENNASLQOKLETREBELKORLLEQAK 1200
DB 1141 EILALQALKEQKAKAESISDKLNDLEKGHAMLENNASLQOKLETREBELKORLLEQAK 1200
QY 1201 LQOQMDLQKNHI FRLTGLQELADRADILKTERSDELYOLENIOVLYSHEKVKMEGTISQ 1260
DB 1201 LQOQMDLQKNHI FRLTGLQELADRADILKTERSDELYOLENIOVLYSHEKVKMEGTISQ 1260
QY 1261 QTKLIDPLQAMQDPQAKKKVPLQYNELKALBEKAKCALEBPLQCTRLIELSAREEA 1320
DB 1261 QTKLIDPLQAMQDPQAKKKVPLQYNELKALBEKAKCALEBPLQCTRLIELSAREEA 1320
QY 1321 AHRKATDHPHPSTATARQQTAMSAIVRSPHQPSAMSLAPSSRRKESSTPEEFSRL 1380

```

Db      1321 AHRATDHPSTPTAATQAQIAMSATVRSPEHQPSAMSLPAPSSRRRESSTPEEFSSRL 1380
Qy      1381 KERHNNIPHRFNVLNRRATKCAVCLDTVHFGROASKLECOVWCHPCKSTCLPATGCL 1440
Db      1381 KERHNNIPHRFNVLNRRATKCAVCLDTVHFGROASKLECOVWCHPCKSTCLPATGCL 1440
Qy      1441 PAEYATHTFEAPCRDKXNSPGIQTEKPESSSLHLEGMKVPRNNKRGQOGMDRKYTVLEGS 1500
Db      1441 PAEYATHTFEAPCRDKXNSPGIQTEKPESSSLHLEGMKVPRNNKRGQOGMDRKYTVLEGS 1500
Qy      1501 KVLVYDNARABAGQPVVEEFELCPDGDVSIHGAGSELANTAADVPYLLKXMSHHT 1560
Db      1501 KVLVYDNARABAGQPVVEEFELCPDGDVSIHGAGSELANTAADVPYLLKXMSHHT 1560
Qy      1561 TCWPGRTLYLAPSPDQKRWYTALESVYAGGRVREKAEADAKLGNLSLLEGGDDRLD 1620
Db      1561 TCWPGRTLYLAPSPDQKRWYTALESVYAGGRVREKAEADAKLGNLSLLEGGDDRLD 1620
Qy      1621 MNCTLPFSDQVVLVGTSEGLVALNVLKNSLTHVPGIYIIKDLKLLMAGEBBA 1680
Db      1621 MNCTLPFSDQVVLVGTSEGLVALNVLKNSLTHVPGIYIIKDLKLLMAGEBBA 1680
Qy      1681 LCLVDVKVKVKGSLAOSHHPADPDISPNIPEAVKGCHLFGAGKTEINGLCICAMPKSVYL 1740
Db      1681 LCLVDVKVKVKGSLAOSHHPADPDISPNIPEAVKGCHLFGAGKTEINGLCICAMPKSVYL 1740
Qy      1741 RYNNELSKYCIKKEIETSEPCSCIFPTVNSILIGNKFEYIDMKOYTLLEFLDKNDHSLA 1800
Db      1741 RYNNELSKYCIKKEIETSEPCSCIFPTVNSILIGNKFEYIDMKOYTLLEFLDKNDHSLA 1800
Qy      1801 PAVFAASNSPVSIVQVNSAGOREBYLLCFHEFGVPVDSYGRSRSTDLLKMSRLPLAFA 1860
Db      1801 PAVFAASNSPVSIVQVNSAGOREBYLLCFHEFGVPVDSYGRSRSTDLLKMSRLPLAFA 1860
Qy      1861 YREPLPTTHNSLEVTIOARSSAGTPARALDIPNRYIGPAISSAATYASGYOKL 1920
Db      1861 YREPLPTTHNSLEVTIOARSSAGTPARALDIPNRYIGPAISSAATYASGYOKL 1920
Qy      1921 RVICCKGNLVKESGTEHHRGPSTSRSSPNKGPPTYNHEITKRVASAPPEGSPHREP 1980
Db      1921 RVICCKGNLVKESGTEHHRGPSTSRSSPNKGPPTYNHEITKRVASAPPEGSPHREP 1980
Qy      1981 STPHRYREGRTLELRDKSPGRPLEREKSPGRILSTRRRSPARLFEDSSRGRLPAGAVRT 2040
Db      1981 STPHRYREGRTLELRDKSPGRPLEREKSPGRILSTRRRSPARLFEDSSRGRLPAGAVRT 2040
Qy      2041 PLSQVNTKVMDOSSV 2054
Db      2041 PLSQVNTKVMDOSSV 2054

RESULT 4
ADJ96610
ID      ADJ96610 standard; protein; 2055 AA.
XX      AC      ADJ96610;
XX      DT      06-MAY-2004 (first entry)
DE      Human citron Rho-interacting kinase CR1K protein SeqID 67.
XX      KW      kinase; human; SNP; single nucleotide polymorphism;
XX      KW      tyrosine protein kinase; serine/threonine protein kinase; PTK; STK;
XX      KW      gene therapy; cancer; immune-related disease; cardiovascular disease;
XX      KW      brain; neuronal associated disease; metabolic; inflammatory disorder;
XX      KW      cytosolic; neuroprotective; immunomodulator; antiinflammatory; enzyme;
XX      KW      citron Rho-interacting kinase; CR1K.
OS      Homo sapiens.
XX      57.
XX      Key      Location/Qualifiers

```

```

FT      Misc-difference 961
FT      /note= "Wild type Ala substituted for Glu by single
FT      nucleotide polymorphism"
XX      PN      WO2004006838-A2.
XX      PD      22-JAN-2004.
XX      PF      15-JUL-2003; 2003WO-US021730.
XX      PR      15-JUL-2002; 2002US-0395632P.
XX      PA      (SUGEN-) SUGEN INC.
XX      PI      Whyte D, Manning G, Caenepeel S;
XX      DR      WPI; 2004-122753/12.
XX      DR      N-PSDB; ADJ96544.
XX      PT      New nucleic acid molecule encoding a kinase polypeptide, useful for
XX      PT      preparing a composition for treating diseases or disorders, e.g., cancer,
XX      PS      or neurological, immunological or inflammatory disorders.
XX      Claim 1; SEQ ID NO 67; 366pp; English.
XX      CC      This invention relates to a novel isolated, enriched or purified nucleic
XX      CC      acid molecule that encodes a kinase polypeptide. Specifically, it relates
XX      CC      to human tyrosine and serine/threonine protein kinases (PTK's and STK's),
XX      CC      as well as protein kinase-like enzymes. The present invention describes
XX      CC      screening methods to identify agonists, antagonists and antibodies that
XX      CC      can be used to modulate the activity or function of the mammalian kinase
XX      CC      enzymes. As such, these compositions can be used for gene therapy
XX      CC      purposes to treat diseases or disorders including cancer, immune-related
XX      CC      diseases, cardiovascular disease, brain or neuronal associated disease,
XX      CC      metabolic and inflammatory disorders. Accordingly, they exhibit
XX      CC      cytosolic, neuroprotective, immunomodulator and anti-inflammatory
XX      CC      activities. This polypeptide sequence is a human kinase protein sequence
XX      CC      of the invention.
XX      SQ      Sequence 2055 AA;
XX      Query Match      99.7%; Score 10464.5; DB 8; Length 2055;
XX      Matched Local Similarity 99.8%; Pred. No. 0;
XX      Matches 2050; Conservative 2; Mismatches 2; Indels 1; Gaps 1;
Qy      1 MLKRYGARNPLDGAAPIASRASRLNLFQKGPPEFTQOQMSPLSREGILDALFVLFE 60
Db      1 MLKRYGARNPLDGAAPIASRASRLNLFQKGPPEFTQOQMSPLSREGILDALFVLFE 60
Qy      61 EGSOPALMKIYGVNSFVRK-YSDTIAELQELQPSAKDFEVRSLVCCGHFAVQVVRKAT 119
Db      61 EGSOPALMKIYGVNSFVRK-YSDTIAELQELQPSAKDFEVRSLVCCGHFAVQVVRKAT 120
Qy      120 GDITAMKVMKKKALLAOQVSEFFEEERNILSRSTSPWIPOLQYAFODGNHLYLWMEYOPG 179
Db      121 GDITAMKVMKKKALLAOQVSEFFEEERNILSRSTSPWIPOLQYAFODGNHLYLWMEYOPG 180
Qy      180 GDLSLNLRYEDQDENLQFYLAELILAVASVHLGMGYVRDIXENILVDRGTGIKLVLD 239
Db      181 GDLSLNLRYEDQDENLQFYLAELILAVASVHLGMGYVRDIXENILVDRGTGIKLVLD 240
Qy      240 FGSAAKNNSNRVNAKLPIGTPTYAPPEVLTVMNGDGKGYGLDGDWMSVGVIAEMLYG 299
Db      241 FGSAAKNNSNRVNAKLPIGTPTYAPPEVLTVMNGDGKGYGLDGDWMSVGVIAEMLYG 300
Qy      300 RSPFAEGTSARTFNINMFORFLKPPDDPKYSSDFLILQSLGQKXRLKFEGLCGHPF 359
Db      301 RSPFAEGTSARTFNINMFORFLKPPDDPKYSSDFLILQSLGQKXRLKFEGLCGHPF 360
Qy      360 FSKIDMNNIRNSPPFPVYTLKSDDDTSNFDPEKNSWVSPPCOLSPSGFSGEELPFVGF 419
Db      361 FSKIDMNNIRNSPPFPVYTLKSDDDTSNFDPEKNSWVSPPCOLSPSGFSGEELPFVGF 420

```


QY 420 SYSKALGILGRSESVVSGLDSPAKTSSMEKLLIKSKELQSDQKCHKMEQEMTLARRV 479
 DB 421 SYSKALGILGRSESVVSGLDSPAKTSSMEKLLIKSKELQSDQKCHKMEQEMTLARRV 480
 QY 480 SEVEAVLSQKEVELKASETQSLREBODLATYITTECSSLKRSLEQARMEVSQEDDALQTL 539
 DB 481 SEVEAVLSQKEVELKASETQSLREBODLATYITTECSSLKRSLEQARMEVSQEDDALQTL 540
 QY 540 HDIRQSKRLQETIKQOEYQAOVEBRLMNOLEBULVSARRSDIYESELRBSRLAAEEF 599
 DB 541 HDIRQSKRLQETIKQOEYQAOVEBRLMNOLEBULVSARRSDIYESELRBSRLAAEEF 600
 QY 600 KRKATECOHKLKAKDOGKPEYGEYAKLEKINABOOLKIQELQELKEXAVKSTAEATL 659
 DB 601 KRKATECOHKLKAKDOGKPEYGEYAKLEKINABOOLKIQELQELKEXAVKSTAEATL 660
 QY 660 QNIRQAKRABERLEKLQNRBDSBGIRKCLVEAEERHSLNKYKRLTEMRERENRLKD 719
 DB 661 QNIRQAKRABERLEKLQNRBDSBGIRKCLVEAEERHSLNKYKRLTEMRERENRLKD 720
 QY 720 DIQTSQOIQOMADKILEEKGREAOVSAOHLVHLKQEQHYEKKIVLDNQIKDILA 779
 DB 721 DIQTSQOIQOMADKILEEKGREAOVSAOHLVHLKQEQHYEKKIVLDNQIKDILA 780
 QY 780 DKETLENNMORHEEBAHEKGLISBOKAMINAMDSKISLBQRIVYELSEBANKLANSSIF 839
 DB 781 DKETLENNMORHEEBAHEKGLISBOKAMINAMDSKISLBQRIVYELSEBANKLANSSIF 840
 QY 840 TORNNKAOBEMISELRQOKFYLETOAGKLEAONRLKLEBOLKISHQSDSDKRLLEETR 899
 DB 841 TORNNKAOBEMISELRQOKFYLETOAGKLEAONRLKLEBOLKISHQSDSDKRLLEETR 900
 QY 900 LREVSLEHEBOYKLEIKROLTELQSLQERESQITLQARALJESOLQAKTELEETAE 959
 DB 901 LREVSLEHEBOYKLEIKROLTELQSLQERESQITLQARALJESOLQAKTELEETAE 960
 QY 960 AABEIQALTARHDEIQRKFDALRNSCTVITDLEBOLNUTEDNALNNONPFLSQDLBA 1019
 DB 961 AABEIQALTARHDEIQRKFDALRNSCTVITDLEBOLNUTEDNALNNONPFLSQDLBA 1020
 QY 1020 SGANDEIVQLRSEVDHLRREITEREMOJTSQKOTMEALKTCTMLEEQVMDLEANDL 1079
 DB 1021 SGANDEIVQLRSEVDHLRREITEREMOJTSQKOTMEALKTCTMLEEQVMDLEANDL 1080
 QY 1080 EKERQWEAMRSVLDGDKSQFECRVNELQRLDTEKQSPARADQRTTESROVVELAVKEXH 1139
 DB 1081 EKERQWEAMRSVLDGDKSQFECRVNELQRLDTEKQSPARADQRTTESROVVELAVKEXH 1140
 QY 1140 AEIILALQALKEOYKLEKASLSIDKNDLEKGTAMLEMMARSLQOKLETREBELKQRLLEBOA 1199
 DB 1141 AEIILALQALKEOYKLEKASLSIDKNDLEKGTAMLEMMARSLQOKLETREBELKQRLLEBOA 1200
 QY 1200 KLOQOMDLOKXHIPLTQGLQALDRADLKTERSDLYOULNOVYSHKEVVMKEGITS 1259
 DB 1201 KLOQOMDLOKXHIPLTQGLQALDRADLKTERSDLYOULNOVYSHKEVVMKEGITS 1260
 QY 1260 QOTKLIDPLQAKMDQPAKKKQVPLQYINELKLALEKAKACALBEALQKTRITELASABE 1319
 DB 1261 QOTKLIDPLQAKMDQPAKKKQVPLQYINELKLALEKAKACALBEALQKTRITELASABE 1320
 QY 1320 AAHRKATDHPHSPATARQOIJAMSATYRSPHOPASMSLAPSSRRKESSTPEESRR 1379
 DB 1321 AAHRKATDHPHSPATARQOIJAMSATYRSPHOPASMSLAPSSRRKESSTPEESRR 1380
 QY 1380 LKERMHNI PHRFNVGLMRAATKAVCLDTVHFGQASKCEQVMCHPKSTCIPATCG 1439
 DB 1381 LKERMHNI PHRFNVGLMRAATKAVCLDTVHFGQASKCEQVMCHPKSTCIPATCG 1440
 QY 1440 LPAEYATHTFEAFCDKKNNSPGLQTKEPSSSLHLEGMKVPANNRKGQGMDRKIVLEG 1499
 DB 1441 LPAEYATHTFEAFCDKKNNSPGLQTKEPSSSLHLEGMKVPANNRKGQGMDRKIVLEG 1500
 QY 1500 SKVLIYDNABABAGORPVEEFBELCLPDGDVSIHGAIVGASELANITAKADVPYILKMESSH 1559

DB 1501 SKVLIYDNABABAGORPVEEFBELCLPDGDVSIHGAIVGASELANITAKADVPYILKMESSH 1560
 QY 1560 TTCWPGRTLYILAPSPDKORWVTALSVVAGRVSRKEXADAKLGNLSLKLKEDDRL 1619
 DB 1561 TTCWPGRTLYILAPSPDKORWVTALSVVAGRVSRKEXADAKLGNLSLKLKEDDRL 1620
 QY 1620 DMNCTLPSPDOVYVGTREGLYALNVLKNSLTHVGCIAVFOIYIITKOLEKLMLAGER 1679
 DB 1621 DMNCTLPSPDOVYVGTREGLYALNVLKNSLTHVGCIAVFOIYIITKOLEKLMLAGER 1680
 QY 1680 ALCLVDVAKVKQSLAQSHLPAQPDISPNIPEAVKGCILFGAGKIENGCLICAMPSKVYI 1739
 DB 1681 ALCLVDVAKVKQSLAQSHLPAQPDISPNIPEAVKGCILFGAGKIENGCLICAMPSKVYI 1740
 QY 1740 LRYNENLSKYCIKKEIETSEPCSIHFTNYSILGTNKFYEIDMKOYTLLEFLDKNDHSL 1799
 DB 1741 LRYNENLSKYCIKKEIETSEPCSIHFTNYSILGTNKFYEIDMKOYTLLEFLDKNDHSL 1800
 QY 1800 APAYFAASSNSFPVSIYVNSAGQREBYLLCFHEFGVFDVSGRRSRDIDLKMSRLPLAF 1859
 DB 1801 APAYFAASSNSFPVSIYVNSAGQREBYLLCFHEFGVFDVSGRRSRDIDLKMSRLPLAF 1860
 QY 1860 AYREBYLFTVTHNSLEVEIEIQRSSAGTPARAYLIDIPNRYLGPALISSGATYLAASYDK 1919
 DB 1861 AYREBYLFTVTHNSLEVEIEIQRSSAGTPARAYLIDIPNRYLGPALISSGATYLAASYDK 1920
 QY 1920 LRVICCKNLVKESGTEHHRGPSTRSSPNRKGPPTYNEHTIKRVAASPAPPEGSHPRE 1979
 DB 1921 LRVICCKNLVKESGTEHHRGPSTRSSPNRKGPPTYNEHTIKRVAASPAPPEGSHPRE 1980
 QY 1980 PSTPHRYEGRTTELRRDQSPGRPLEREXSPGRITLSTRERSPARLPESSRGRLPAGAVR 2039
 DB 1981 PSTPHRYEGRTTELRRDQSPGRPLEREXSPGRITLSTRERSPARLPESSRGRLPAGAVR 2040
 QY 2040 TPLSOVNVKWDQSSV 2054
 DB 2041 TPLSOVNVKWDQSSV 2055

RESULT 5
 ABG78363
 ID ABG78363 standard; protein, 2053 AA.
 XX AC ABG78363;
 XX AC XX
 XX DT 15-NOV-2002 (first entry)
 XX DE RHO/RAC-interacting citron kinase-like human protein, designated NOV3b.
 XX
 XX Human; NOV; NOVX; NOVX-associated disorder; cardiomyopathy;
 XX atherosclerosis; diabetes; cell signaling; metabolic pathway;
 XX cellular receptor; downstream effector; cancer; gene therapy;
 XX hypertension; congenital heart defect; aortic stenosis; obesity;
 XX infectious disease; anorexia; Alzheimer's disease; Parkinson's disease;
 XX neurodegenerative disorder; haemophilia; dyslipidemia; vaccine;
 XX haematopoietic disease; scleroderma; fertility; immunogen;
 XX idiopathic thrombocytopenic purpura; graft versus host disease;
 XX Crohn's disease; multiple sclerosis; cirrhosis; autoimmune disease;
 XX systemic lupus erythematosus; asthma; arthritis; psoriasis; allergy;
 XX stroke; anxiety; Leach-Nyhan syndrome; schizophrenia; cerebellar ataxia;
 XX pain; alcoholism; transgenic.
 OS Homo sapiens.
 XX
 XX MO200226826-A2.
 XX
 XX PD 04-APR-2002.
 XX
 XX PF 27-SEP-2001; 2001WO-US042336.
 XX
 XX PR 27-SEP-2000; 2000US-0235631P.
 XX
 XX PR 27-SEP-2000; 2000US-0235633P.

PR 27-SEP-2000; 2000US-0235808P.
 PR 27-SEP-2000; 2000US-0236064P.
 PR 27-SEP-2000; 2000US-0236065P.
 PR 27-SEP-2000; 2000US-0236066P.
 PR 28-SEP-2000; 2000US-0236135P.
 PR 03-OCT-2000; 2000US-0237434P.
 PR 05-OCT-2000; 2000US-0238321P.
 PR 06-OCT-2000; 2000US-0238396P.
 PR 06-OCT-2000; 2000US-0238399P.
 PR 16-MAR-2001; 2001US-0276667P.
 PR 31-MAY-2001; 2001US-0294823P.
 PR 12-JUL-2001; 2001US-0304868P.
 PR 26-SEP-2001; 2001US-00235631.
 XX
 PA (CURA-) CURAGEN CORP.
 XX
 PI Gerlach VL, MacDougall JR, Smithson G, Millet I, Stone D;
 PI Gunther E, Ellerman K, Grose WM, Alsbrook JP, Leplay DM;
 PI Burgess CE, Padigara M, Kekuda R, Spytek KA, Leach MD, Shinkens RA;
 XX
 DR WPI; 2002-499860/53.
 DR N-PSDB; ABS63436.
 XX

Novel isolated NOVX polypeptides and polynucleotides homologous to attractin, plexin, papin-like family of proteins, useful for treating atherosclerosis, diabetes, cancer, Alzheimer's disease, hemophilia and stroke.

Claim 1; Page 44-45; 308bp; English.

The invention discloses the isolated human polypeptides, and polynucleotides encoding them, that have been designated NOVX. The polypeptides, polynucleotides and antibodies are useful in treating or preventing a NOVX-associated disorder which is cardiomyopathy, atherosclerosis and diabetes in a human, where the disorder is related to cell signal processing and metabolic pathway modulation. They can also be used in determining the presence of, or predisposition to, a disease associated with altered levels of the polypeptides and polynucleotides of any one of the 13 sequences (NOV1-NOV8), for raising antibodies, for identifying an agent that binds to, or that modulates the expression or activity of the polypeptide, for identifying an agent which is cellular receptor or downstream effector, for treating or preventing a NOVX-associated disorder and as a pharmaceutical composition comprising the polypeptide, polynucleotide or the antibody. The polypeptides and polynucleotides are useful in diagnostic applications (e.g. as a marker for cancerous cells or tissue types) where their amounts are assessed, or for the manufacture of a medicament (e.g. gene therapy) for treating or preventing disorders or syndromes such as hypertension, congenital heart defects, aortic stenosis, obesity, infectious disease, anorexia, cancer, Alzheimer's disease, Parkinson's disorders, neurodegenerative disorders, haemophilia, dyslipidemias, haematopoietic diseases, scleroderma, fertility, idiopathic thrombocytopenic purpura, graft versus host diseases, Crohn's disease, multiple sclerosis, cirrhosis, autoimmune disease, systemic lupus erythematosus, asthma, arthritis, psoriasis, allergy, stroke, anxiety, Lesch-Nyhan syndrome, schizophrenia, cerebellar ataxia, pain and alcoholism. They may also be used as immunogens to produce antibodies specific for the invention, and as vaccines. Transgenic cells containing a NOVX expressing construct are useful to produce non-human transgenic animals for studying the function and/or activity of the NOVX proteins and for identifying and/or evaluating modulators of NOVX protein activity. Transgenic cells containing a NOVX expressing construct are useful to produce non-human transgenic animals for identifying and/or evaluating modulators of NOVX protein activity. Transgenic cells containing a NOVX expressing construct are useful to produce non-human transgenic animals for studying the function and/or activity of the NOVX proteins and for identifying and/or evaluating modulators of NOVX protein activity. The sequences presented in ABG78359-ABG78371 are the human NOV1-NOV8 proteins

XX Sequence 2053 AA;

Query Match 99.7%; Score 10458.5; DB 5; Length 2053;
 Best Local Similarity 99.8%; Pred. No. 0;
 Matches 2049; Conservative 2; Mismatches 2; Indels 1; Gaps 1;

1 MLKFRYGARNPIDAGAEPIASRASRLNLFQGGKPPFTQQQMSPISREGIIDALFVLFE 60

Db 1 MLKFRYGARNPIDAGAEPIASRASRLNLFQGGKPPFTQQQMSPISREGIIDALFVLFE 60
 QY 61 ECSPALMKIKHVSNPVAKYSDDTIAEIOELPSAKDFEVRSLVGGHFAVQVYREKATG 120
 Db 61 ECSPALMKIKHVSNPVAKYSDDTIAEIOELPSAKDFEVRSLVGGHFAVQVYREKATG 120
 QY 121 DIYAMKWKKKALLAOEVSFFEEERNLSTSPWIPQLOYAPQDKNHLVWVEYQGG 180
 Db 121 DIYAMKWKKKALLAOEVSFFEEERNLSTSPWIPQLOYAPQDKNHLVWVEYQGG 180
 QY 181 DLSILNRKEDQDENLQFIYLAELILAVSHVLMGVHARDIKPENILVDRTHIKLVDF 240
 Db 181 DLSILNRKEDQDENLQFIYLAELILAVSHVLMGVHARDIKPENILVDRTHIKLVDF 240
 QY 241 GSAAKMNSNMVNAKLPIGPDYNAPEVLTVANGDGKGTGGLDCDMSVGIAYIMYIGR 300
 Db 241 GSAAKMNSNMVNAKLPIGPDYNAPEVLTVANGDGKGTGGLDCDMSVGIAYIMYIGR 300
 QY 301 SPFAEGTSARTFNINMFQRFKPPDPKVSDFLDLIQSLCCQKRLKEGGLCCHPFF 359
 Db 301 SPFAEGTSARTFNINMFQRFKPPDPKVSDFLDLIQSLCCQKRLKEGGLCCHPFF 359
 QY 361 SKIDNNIRNSPPFPVPTLKSDDDTSPNDEPKNSWSSPCQSPSGSGEELPFVGF 420
 Db 361 SKIDNNIRNSPPFPVPTLKSDDDTSPNDEPKNSWSSPCQSPSGSGEELPFVGF 420
 QY 420 YSKALGILGRSESVVSGIDSPAKTSMEKLLIKSKELQSDQKCHKEQETRLHRRVS 479
 Db 420 YSKALGILGRSESVVSGIDSPAKTSMEKLLIKSKELQSDQKCHKEQETRLHRRVS 479
 QY 481 EVEAVLSQKVELKASEFQRBILRODLATYTTGSSLSKRSLEQARMEVSQDDKLOLH 540
 Db 481 EVEAVLSQKVELKASEFQRBILRODLATYTTGSSLSKRSLEQARMEVSQDDKLOLH 540
 QY 480 EVEAVLSQKVELKASEFQRBILRODLATYTTGSSLSKRSLEQARMEVSQDDKLOLH 539
 Db 480 EVEAVLSQKVELKASEFQRBILRODLATYTTGSSLSKRSLEQARMEVSQDDKLOLH 539
 QY 541 DIRESRKLOETIKQEOYAOVEKRLMNOLEBDIVSARRSDIYESLSRLAAEFK 600
 Db 541 DIRESRKLOETIKQEOYAOVEKRLMNOLEBDIVSARRSDIYESLSRLAAEFK 600
 QY 540 DIRESRKLOETIKQEOYAOVEKRLMNOLEBDIVSARRSDIYESLSRLAAEFK 599
 Db 540 DIRESRKLOETIKQEOYAOVEKRLMNOLEBDIVSARRSDIYESLSRLAAEFK 599
 QY 601 RKATECOHKLKANDQKPEVEGYAKLEKINAEOQLKIOELQEKLEKAVKASTATELILQ 660
 Db 601 RKATECOHKLKANDQKPEVEGYAKLEKINAEOQLKIOELQEKLEKAVKASTATELILQ 660
 QY 660 NROKAPAEERLEKLOQREDSSEGRKLVABERRRSLNKYKRLTMRERRRLDD 720
 Db 660 NROKAPAEERLEKLOQREDSSEGRKLVABERRRSLNKYKRLTMRERRRLDD 720
 QY 721 IOTSSQOIQWADKILIEEKHREAOVSAHLEVLKQEOHYEEKIVLDNQIKKDLAD 780
 Db 721 IOTSSQOIQWADKILIEEKHREAOVSAHLEVLKQEOHYEEKIVLDNQIKKDLAD 780
 QY 780 KETLENMQRHEEAEHEKGLISSEKAMINAMDSKIRSLBQRIYELSEANKLAANSILFT 840
 Db 780 KETLENMQRHEEAEHEKGLISSEKAMINAMDSKIRSLBQRIYELSEANKLAANSILFT 840
 QY 841 QNNKAPQEWMLSELRQKQFYLETOAGKLEPAQNRKLEBQLEKISHODHDKRLLLETRL 900
 Db 841 QNNKAPQEWMLSELRQKQFYLETOAGKLEPAQNRKLEBQLEKISHODHDKRLLLETRL 900
 QY 840 QNNKAPQEWMLSELRQKQFYLETOAGKLEPAQNRKLEBQLEKISHODHDKRLLLETRL 899
 Db 840 QNNKAPQEWMLSELRQKQFYLETOAGKLEPAQNRKLEBQLEKISHODHDKRLLLETRL 899
 QY 901 REVSLHEEQLKLEKQLELOLSLOERESQTLQAARALLESOLROAKTELEETTAB 960
 Db 901 REVSLHEEQLKLEKQLELOLSLOERESQTLQAARALLESOLROAKTELEETTAB 960
 QY 900 REVSLHEEQLKLEKQLELOLSLOERESQTLQAARALLESOLROAKTELEETTAB 959
 Db 900 REVSLHEEQLKLEKQLELOLSLOERESQTLQAARALLESOLROAKTELEETTAB 959
 QY 961 EEEIQTALAHNDEIQKFDALRNSCTVITDLEBQNLQTEDNAELNNQFYLKQLDAS 1020
 Db 961 EEEIQTALAHNDEIQKFDALRNSCTVITDLEBQNLQTEDNAELNNQFYLKQLDAS 1020
 QY 960 EEEIQTALAHNDEIQKFDALRNSCTVITDLEBQNLQTEDNAELNNQFYLKQLDAS 1019
 Db 960 EEEIQTALAHNDEIQKFDALRNSCTVITDLEBQNLQTEDNAELNNQFYLKQLDAS 1019
 QY 1021 GANDEIVQLRSEVDHLREITEREQMUTSQQVTEALKTCTMAEEQVMDLALNDELLE 1080
 Db 1021 GANDEIVQLRSEVDHLREITEREQMUTSQQVTEALKTCTMAEEQVMDLALNDELLE 1080
 QY 1020 GANDEIVQLRSEVDHLREITEREQMUTSQQVTEALKTCTMAEEQVMDLALNDELLE 1079
 Db 1020 GANDEIVQLRSEVDHLREITEREQMUTSQQVTEALKTCTMAEEQVMDLALNDELLE 1079
 QY 1081 KERQWEARSTVLGPEKSGFECGRVLEQMLDTEKQSRARAORTRESQVVELAVKEKA 1140
 Db 1081 KERQWEARSTVLGPEKSGFECGRVLEQMLDTEKQSRARAORTRESQVVELAVKEKA 1140

Db 1080 KERQWENMSVLGDEKSOFECEVRRELQRMIDTEKOSRARDQRTESRQVVELAVKENKA 1139
Qy 1141 EILALQOALKEQKLKAESLSDKLNDEKKAEMENNARSLOQKLETERELKQELLEBOAK 1200
Db 1140 EILALQOALKEQKLKAESLSDKLNDEKKAEMENNARSLOQKLETERELKQELLEBOAK 1199
Qy 1201 LQOQMDLOKXNHFRLTQGLQALDRADDLKTERSDELEYQLENIQVLYSHKXKMEGTISQ 1260
Db 1200 LQOQMDLOKXNHFRLTQGLQALDRADDLKTERSDELEYQLENIQVLYSHKXKMEGTISQ 1259
Qy 1261 QTKLIDFLQAKMOPAKKKKVPLOYNELKALEKXKARCALEBEALQKTEILRSAREEA 1320
Db 1260 QTKLIDFLQAKMOPAKKKKVPLOYNELKALEKXKARCALEBEALQKTEILRSAREEA 1319
Qy 1321 AHRKATDHPHSTPATARQOIAMSAIVRSPEHOPSMASLIAPSSRKESSTPEERSRL 1380
Db 1320 AHRKATDHPHSTPATARQOIAMSAIVRSPEHOPSMASLIAPSSRKESSTPEERSRL 1379
Qy 1381 KERHMHNIPIHRENVGLNMRATKCAVCLDTHVFGROASKLECOVMCHPKSTCLPATCGL 1440
Db 1380 KERHMHNIPIHRENVGLNMRATKCAVCLDTHVFGROASKLECOVMCHPKSTCLPATCGL 1439
Qy 1441 PARYATHFTEAFRCDDKNSPGLQTKEPSSSLHLEGMKVPRNNKRGQGDRTKTYLVEGS 1500
Db 1440 PARYATHFTEAFRCDDKNSPGLQTKEPSSSLHLEGMKVPRNNKRGQGDRTKTYLVEGS 1499
Qy 1501 KVLIIYDNEAREAGORPVEEFELCLPDGVSIGHVAGASELANTAKADVPIYILKXESHPT 1560
Db 1500 KVLIIYDNEAREAGORPVEEFELCLPDGVSIGHVAGASELANTAKADVPIYILKXESHPT 1559
Qy 1561 TCMGRTLYLLAPSFPDKQKWVTALLESVAGRSREKADAKLLGNSILKLEGGDRLD 1620
Db 1560 TCMGRTLYLLAPSFPDKQKWVTALLESVAGRSREKADAKLLGNSILKLEGGDRLD 1619
Qy 1621 MNCITLPPSDQVNVGTEBEGYALNVKNSLTHVPGTGAHVQIYIILKOLEKLMTAGERA 1680
Db 1620 MNCITLPPSDQVNVGTEBEGYALNVKNSLTHVPGTGAHVQIYIILKOLEKLMTAGERA 1679
Qy 1681 LCLVDYKVKQSLAOSHLPAPDISPNIPEAVKGCILFGAGKTENGLCIQAMPSSKVIL 1740
Db 1680 LCLVDYKVKQSLAOSHLPAPDISPNIPEAVKGCILFGAGKTENGLCIQAMPSSKVIL 1739
Qy 1741 RYNNENSKYCIKREIETSEPCSHFTNYSILIGTNKFEYIDMKQYTLSEFLDKNDHSLA 1800
Db 1740 RYNNENSKYCIKREIETSEPCSHFTNYSILIGTNKFEYIDMKQYTLSEFLDKNDHSLA 1799
Qy 1801 PAVPAASNSFPVSIYOVNSAGOREEYLLCFHEGCVFVDSYGRSRITDDIKWSRLPLAFA 1860
Db 1800 PAVPAASNSFPVSIYOVNSAGOREEYLLCFHEGCVFVDSYGRSRITDDIKWSRLPLAFA 1859
Qy 1861 YREBPVLFVTHPNSLEVEIEIOARSAGTARAYLDIPNRYLGPATISSGATYILASSYODKL 1920
Db 1860 YREBPVLFVTHPNSLEVEIEIOARSAGTARAYLDIPNRYLGPATISSGATYILASSYODKL 1919
Qy 1921 RVLCCKGNLVKESGTEHHRGSPSTRSSPNKGPPTVNEHTTKRVASSPAPPEGSHPREP 1980
Db 1920 RVLCCKGNLVKESGTEHHRGSPSTRSSPNKGPPTVNEHTTKRVASSPAPPEGSHPREP 1979
Qy 1981 STPHRYEGRTTELRRDKSPGRPLERKSPGRITLSTRERSPARLFEDSSRGRLPAQAVRT 2040
Db 1980 STPHRYEGRTTELRRDKSPGRPLERKSPGRITLSTRERSPARLFEDSSRGRLPAQAVRT 2039
Qy 2041 PLSOVNVKWDQSSV 2054
Db 2040 PLSOVNVKWDQSSV 2053

RESULT 6

ADA05642

ID ADA05642 standard; protein: 2053 AA.

XX AC ADA05642;

XX

XX

DT 06-NOV-2003 (first entry)
DE Human NOVA protein SEQ ID NO:2.
KW human; NOVA; antidiabetic; anorectic; antibacterial; virocidic;
KW immunomodulator; cytostatic; neuroprotective;
KW metapartinsin; antidiabetic; gene therapy; human disease;
KW metabolic disorder; diabetes; obesity; infection; cachexia; cancer;
KW neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;
KW immune disorder; haematopoietic disorder; dyslipidaemia.
XX Homo sapiens.
XX WO2003029424-A2.
XX 10-APR-2003.
XX 02-OCT-2002; 2002WO-US031373.
XX 02-OCT-2001; 2001US-0326483P.
XX 05-OCT-2001; 2001US-0327435P.
XX 09-OCT-2001; 2001US-0327449P.
XX 09-OCT-2001; 2001US-0327917P.
XX 09-OCT-2001; 2001US-0328029P.
XX 09-OCT-2001; 2001US-0328044P.
XX 09-OCT-2001; 2001US-0328056P.
XX 12-OCT-2001; 2001US-0328849P.
XX 15-OCT-2001; 2001US-0329414P.
XX 17-OCT-2001; 2001US-0330142P.
XX 18-OCT-2001; 2001US-0330309P.
XX 22-OCT-2001; 2001US-0341058P.
XX 24-OCT-2001; 2001US-0339266P.
XX 24-OCT-2001; 2001US-0343629P.
XX 29-OCT-2001; 2001US-0349575P.
XX 01-NOV-2001; 2001US-0346357P.
XX 17-APR-2002; 2002US-0373260P.
XX 19-APR-2002; 2002US-0373815P.
XX 19-APR-2002; 2002US-0373817P.
XX 19-APR-2002; 2002US-0373826P.
XX 19-APR-2002; 2002US-0373884P.
XX 22-APR-2002; 2002US-0374977P.
XX 16-MAY-2002; 2002US-0381037P.
XX 16-MAY-2002; 2002US-0381038P.
XX 16-MAY-2002; 2002US-0381042P.
XX 17-MAY-2002; 2002US-0381642P.
XX 28-MAY-2002; 2002US-0383656P.
XX 29-MAY-2002; 2002US-0383831P.
XX 25-JUN-2002; 2002US-0391335P.
XX 01-OCT-2002; 2002US-00262511.
XX (CURA-) CURAGEN CORP.
XX Smtneon G, Millet I, Peyman JA, Kekuda R, Ju J, Li L, Guo X;
PI Paturajan M, Spytek KA, Bainger SR, Ellerman K, Malvankar UM;
PI O T, Gorman L, Zetunuen BD, Anderson DW, Zhong M, Catterton E;
PI Ji W, Miller CE, Raselli L, Stone DJ, Pena CE, Shenoy SC;
PI Shinkets RA, Rothenberg ME, Leach MD, Agee ML, Bergins C, Dippio VA;
PI Eisen AJ, Gangoli EA, Rieger DK, Spaderna SK;
XX WPI; 2003-381626/36.
XX N-PSDB; ADA05641.
XX New NOVA polypeptides and nucleic acids, useful for diagnosing,
PT preventing or treating NOVA-associated disorders, e.g. diabetes, obesity,
PT cancer or dyslipidemia, and in chromosome mapping, tissue typing or
PT pharmacogenomics.
XX Claim 1; Page 99-100; 586pp; English.
XX The present invention describes NOVA proteins, where X can be 1 to 55
CC (e.g. NOVA). Also described: (1) a composition comprising a polypeptide
CC described above and a carrier; (2) a kit comprising, in one or more
CC containers, the composition described above; (3) an isolated nucleic acid

CC molecule which encodes a NOVX protein of the invention; (4) a vector
CC comprising the nucleic acid molecule described above; (5) a cell
CC comprising the above vector; (6) an antibody that immunospecifically
CC binds to the polypeptide described above; (7) methods for determining the
CC presence or amount of the above polypeptide or nucleic acid molecule in a
CC sample; (8) methods for determining the presence of or predisposition to
CC a disease associated with altered levels of expression of the above
CC polypeptide or nucleic acid molecule in a first mammalian subject; (9) a
CC method of identifying an agent that binds to the polypeptide described
CC above; (10) a method for identifying a potential therapeutic agent for
CC use in treating a pathology that is related to an aberrant expression or
CC aberrant physiological interactions of the polypeptide; (11) a method of
CC screening for a modulator of activity or of latency or predisposition to
CC a pathology associated with the polypeptide; (12) a method for modulating
CC the activity of the polypeptide described above; (13) methods of treating
CC or preventing a pathology associated with the above polypeptide in a
CC mammal; and (14) a method for producing the above polypeptide. NOVX
CC sequences have antidiabetic, anorectic, antibacterial, virocidic,
CC immunomodulator, cytostatic, nootropic, neuroprotective, antiparkinsonian
CC and antilipemic activities, and can be used in gene therapy. The
CC polypeptide is useful in manufacturing a medicament for treating a
CC syndrome associated with a human disease. The polypeptide or the nucleic
CC acid molecule may be used to diagnose, treat or prevent metabolic
CC disorders such as diabetes or obesity, infections, cachexia, cancer,
CC neurodegenerative disorders such as Alzheimer's disease or Parkinson's
CC disease, immune disorders, hematopoietic disorders and various
CC dyslipidaemias. The nucleic acids can also be used as hybridisation
CC probes, in chromosome mapping, tissue typing, preventive medicine and
CC pharmacogenomics. The present sequence represents a human NOVX from the
CC present invention.

XX Sequence 2053 AA;

SQ Query Match 99.7%; Score 10458.5; DB 6; Length 2053;

Best Local Similarity 99.8%; Pred. No. 0;

Matches 2049; Conservative 2; Mismatches 2; Indels 1; Gaps 1;

QY 1 MKFKYGANNPLDAGAAEPPIASASRLNLPFGCKPPPTQOQMSPLSRBGILDALFVLE 60
DB 1 MKFKYGANNPLDAGAAEPPIASASRLNLPFGCKPPPTQOQMSPLSRBGILDALFVLE 60
QY 61 ECSOPALMKIKHVSNFVKYSDTIAELOPSAKDFEVRSLVGCCHPAVQVYVRKAG 120
DB 61 ECSOPALMKIKHVSNFVKYSDTIAELOPSAKDFEVRSLVGCCHPAVQVYVRKAG 120
QY 121 DIYAMKWMKKALLAQEQVSPFEERNILSRSTSPWIPOLQYAFODKNLILYVMEYQPG 180
DB 121 DIYAMKWMKKALLAQEQVSPFEERNILSRSTSPWIPOLQYAFODKNLILYVMEYQPG 180
QY 181 DLISLNRREDQDENLLOFYLAELILAVSHLMGYVHRDIKPNILVDRGTGHIKLVDF 240
DB 181 DLISLNRREDQDENLLOFYLAELILAVSHLMGYVHRDIKPNILVDRGTGHIKLVDF 240
QY 241 GSAAMKNSKMNNAKLPIGTPDYMAPEVLTNNNGDGKGYGLDCMWSGVLAEMTYR 300
DB 241 GSAAMKNSKMNNAKLPIGTPDYMAPEVLTNNNGDGKGYGLDCMWSGVLAEMTYR 300
QY 301 SPFAGTSARTNNINNFQFLKFPDDPKVSSDFDLLOSILCGQKRLKFEGLCCHPFF 360
DB 301 SPFAGTSARTNNINNFQFLKFPDDPKVSSDFDLLOSILCGQKRLKFEGLCCHPFF 360
QY 360 SKIDMNNIRNSPPPVFTLKSDDDTSNDFEPKNSWSSPCOLSPGSGEELPVGVS 420
DB 360 SKIDMNNIRNSPPPVFTLKSDDDTSNDFEPKNSWSSPCOLSPGSGEELPVGVS 420
QY 421 YSKALGIIIGRSSSVSGLDSPAKTSSMEKKLILKSELEODSODCKHMEQENTRIHRRVS 480
DB 421 YSKALGIIIGRSSSVSGLDSPAKTSSMEKKLILKSELEODSODCKHMEQENTRIHRRVS 480
QY 481 EYEAVTLQKEVELKASETQSLLEODLATYITEGSSLKRSLEQAMEVSOEDKALQILH 540
DB 481 EYEAVTLQKEVELKASETQSLLEODLATYITEGSSLKRSLEQAMEVSOEDKALQILH 540
QY 540 EYEAVTLQKEVELKASETQSLLEODLATYITEGSSLKRSLEQAMEVSOEDKALQILH 539

QY 541 DIREQSRKLOEITEGEYQAOVEEMRLMMNOLEEDLYSARRSDLYSESLAAEEFK 600
DB 540 DIREQSRKLOEITEGEYQAOVEEMRLMMNOLEEDLYSARRSDLYSESLAAEEFK 599
QY 601 RKATECOHKLKAKDQKPEVGYALKEKINNAQOLKIOLQKLEKAVASTAETLLO 660
DB 600 RKATECOHKLKAKDQKPEVGYALKEKINNAQOLKIOLQKLEKAVASTAETLLO 659
QY 661 NIRAQERARELEKLONRDSESGIRKCLVEAEERHSLLENKVKLETMERREBNRKD 720
DB 660 NIRAQERARELEKLONRDSESGIRKCLVEAEERHSLLENKVKLETMERREBNRKD 719
QY 721 IQRSQOIQOMADKIILEBKHEAQAQSHLEVLKQKQHYEEKIKVLDNQIKODLAD 780
DB 720 IQRSQOIQOMADKIILEBKHEAQAQSHLEVLKQKQHYEEKIKVLDNQIKODLAD 779
QY 781 KETLEMMORHEEAHEKGIKISEQAMINAMDSKTRSORIVERSEANKLAANSLEFT 840
DB 780 KETLEMMORHEEAHEKGIKISEQAMINAMDSKTRSORIVERSEANKLAANSLEFT 839
QY 841 ORNMKAQEMISELRQOKFYLETQAGKLEAQRKLEBOLEKISHODSPKNRLLLEETRL 900
DB 840 ORNMKAQEMISELRQOKFYLETQAGKLEAQRKLEBOLEKISHODSPKNRLLLEETRL 899
QY 901 REVSLEHEBQKLELKQULTLOLSIQERESQULTALQAPALRESQURQAKTELETTAA 960
DB 900 REVSLEHEBQKLELKQULTLOLSIQERESQULTALQAPALRESQURQAKTELETTAA 959
QY 961 EBEIQALTARBEIQKPKPALNSCTVITDLEQONLVEDNAELNNQNPYLSKQIDEAS 1020
DB 960 EBEIQALTARBEIQKPKPALNSCTVITDLEQONLVEDNAELNNQNPYLSKQIDEAS 1019
QY 1021 GANDEIVQURSEVDHLRREITEREMQILSQKOTMEALKTCTMLEBQVWDLLEALNDEL 1080
DB 1020 GANDEIVQURSEVDHLRREITEREMQILSQKOTMEALKTCTMLEBQVWDLLEALNDEL 1079
QY 1081 KERQWEAMRSVLGDEKSOPECEVRBELQRMLDTEKOSRAPADORTESROVVELAVEHKA 1140
DB 1080 KERQWEAMRSVLGDEKSOPECEVRBELQRMLDTEKOSRAPADORTESROVVELAVEHKA 1139
QY 1141 ETLALQALKEOKLAEESLSDKLEKKALEMMANSLQOKLETERELKORLLEBOAK 1200
DB 1140 ETLALQALKEOKLAEESLSDKLEKKALEMMANSLQOKLETERELKORLLEBOAK 1199
QY 1201 LOQOMDLOKNHIFRLTQGLQELADRADILKTERSDLEYOLENIQVLYSHEKYMESTISQ 1260
DB 1200 LOQOMDLOKNHIFRLTQGLQELADRADILKTERSDLEYOLENIQVLYSHEKYMESTISQ 1259
QY 1261 QTKLIDFLQAKMDQPAKKKVPLOYNELKALEKEKARCALIEBALQKTRIELRSAREBA 1320
DB 1260 QTKLIDFLQAKMDQPAKKKVPLOYNELKALEKEKARCALIEBALQKTRIELRSAREBA 1319
QY 1321 AHRKATDHPHSTPTAYARQOIAMSAIVRSPHQPAMSLAPPSSRRKXESTPPEFSRRL 1380
DB 1320 AHRKATDHPHSTPTAYARQOIAMSAIVRSPHQPAMSLAPPSSRRKXESTPPEFSRRL 1379
QY 1381 KERMEHNTPHRPNVGIANRATKCAVCLPTVFHGRQASKLEQOWCHPKCSTCLPATGCL 1440
DB 1380 KERMEHNTPHRPNVGIANRATKCAVCLPTVFHGRQASKLEQOWCHPKCSTCLPATGCL 1439
QY 1441 PAEYATHTFEAFCRDKMNSPGIQTEPESSLHLBGMKVPRNNKRGQOGMDRKYTVLBS 1500
DB 1440 PAEYATHTFEAFCRDKMNSPGIQTEPESSLHLBGMKVPRNNKRGQOGMDRKYTVLBS 1499
QY 1501 KYLTYDNEARAGQRPVEFEELCLPDGVYSIHGAVGASELANTAKADVPYILKMSHPHT 1560
DB 1500 KYLTYDNEARAGQRPVEFEELCLPDGVYSIHGAVGASELANTAKADVPYILKMSHPHT 1559
QY 1561 TCMGRTLYTLAPSPDDQORWYTLAESVAVAGRSRREARADAKLIGNSLLTLEGDDDL 1620
DB 1560 TCMGRTLYTLAPSPDDQORWYTLAESVAVAGRSRREARADAKLIGNSLLTLEGDDDL 1619
QY 1621 MNCITLPPSDQVVLVGTBGLYALNVLKNSLTHVPGI GAVPQIYIHKDLEKLMTAGEBRA 1680

Db 1620 MNCTLPEDQVIVLQTEBGLVALANVLKNSLTHVPEIGAVFOIYIKDKLEKLMINGERRA 1679
Qy 1681 LCLVDVKKVKQSLAQSHPADPDISPNI FEAVKGCHFGAGKIENGLCICAMPBKVIL 1740
Db 1680 LCLVDVKKVKQSLAQSHPADPDISPNI FEAVKGCHFGAGKIENGLCICAMPBKVIL 1739
Qy 1741 RYNNENLSKYCIKKEIETSEPCSHFTNTYSILIGNKVEYEDMKQYTLBEPFLDKNDHSLA 1800
Db 1740 RYNNENLSKYCIKKEIETSEPCSHFTNTYSILIGNKVEYEDMKQYTLBEPFLDKNDHSLA 1799
Qy 1801 PAVEAASNSPSPVSIQVNSAGOREEYLICFHEGVPFVDSYGRSRRTDMLKMSRLPLAFA 1860
Db 1800 PAVEAASNSPSPVSIQVNSAGOREEYLICFHEGVPFVDSYGRSRRTDMLKMSRLPLAFA 1859
Qy 1861 YREPYLFTVTHNSLEVIIRIOARSSAGTPARAVIDIPNRYLQPAISSGAIYLAASSYQDKL 1920
Db 1860 YREPYLFTVTHNSLEVIIRIOARSSAGTPARAVIDIPNRYLQPAISSGAIYLAASSYQDKL 1919
Qy 1921 RVLCCKGNLVKESGTEHHRGPSTSSPNKGGPPYNNHITKRVASSPAPBEGSPHREP 1980
Db 1920 RVLCCKGNLVKESGTEHHRGPSTSSPNKGGPPYNNHITKRVASSPAPBEGSPHREP 1979
Qy 1981 STPHRYBGRTELRRDKSPGRPLERKSPGRILSTRRRSPARLPEDSSRGRLPAQAVT 2040
Db 1980 STPHRYBGRTELRRDKSPGRPLERKSPGRILSTRRRSPARLPEDSSRGRLPAQAVT 2039
Qy 2041 PLSQVNMKVMDOSSV 2054
Db 2040 PLSQVNMKVMDOSSV 2053

RESULT 7
AAU03501

ID AAU03501 standard; protein; 2053 AA.

XX AC AAU03501;
XX DT 12-SEP-2001 (first entry)
XX DE Human protein kinase #1.
XX KM Human, protein kinase; PTK; STK; cancer; cardiovascular disease;
KM metabolic disorder; immune related disease; neurological disorder;
KM neurodegenerative disorder; inflammatory disorder; infectious disease;
KM reproductive disorder.
XX OS Homo sapiens.
XX PN M0200138503-A2.
XX PD 31-MAY-2001.
XX PF 22-NOV-2000; 2000M0-US032085.
XX PR 24-NOV-1999; 99US-0167482P.
XX PA (SUGC-) SUGEN INC.
XX PI Pijman GD, Whyte D, Manning G, Sudarsanam S, Martinez R;
PI Flanagan P, Clary D;
XX MPI; 2001-343950/36.
XX DR N-PSDB; AAS06701.
XX PT Nucleic acid encoding human kinase polypeptides, useful for preventing
PT diagnosing and/or treating e.g. cancer, immune, cardiovascular and
PT neuronal-associated diseases, and microbial infections.
XX PS Claim 7; Fig 2; 433pp; English.
XX CC AAU03501-AAU03557 represent novel human protein kinases #1-57. The novel
CC protein kinases have been identified as members of the tyrosine or

CC serine/threonine kinase (PK and STK) families. The polynucleotides
CC encoding protein kinases and the polypeptides may be used in the
CC prevention, diagnosis and treatment of diseases associated with
CC inappropriate kinase expression. For example, they may be used to treat
CC cancers (especially cancers of haematopoietic origin), cardiovascular
CC disease (e.g. atherosclerosis), metabolic disorders (e.g. diabetes),
CC immune related diseases (e.g. rheumatoid arthritis), neurological
CC disorders (e.g. schizophrenia), neurodegenerative disorders (e.g.
CC Parkinson's disease), inflammatory disorders (e.g. asthma), infectious
CC disease (e.g. HIV) and reproductive disorders (e.g. infertility).
CC Additionally, polynucleotides encoding protein kinases may be used for
CC gene therapy and as DNA probes in diagnostic assays. The protein kinase
CC polypeptides may be used as antigens in the production of antibodies
CC against the protein kinases and in assays to identify modulators of
CC protein kinase expression and activity
SQ Sequence 2053 AA;
Query Match 99.4%; Score 10431.5; DB 4; Length 2053;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 2045; Conservative 2; Mismatches 5; Indels 1; Gaps 1;
Qy 1 MLKFRYGNPLDGAEPISRASRLNLFQGGKPPFTQOQMSPLSEGLIDLALFYVE 60
Db 1 MLKFRYGNPLDGAEPISRASRLNLFQGGKPPFTQOQMSPLSEGLIDLALFYVE 60
Qy 61 ECSQPALMKIKGVNPFVRK-YSDTIAELOLQPSAKDFEVSLVCGHFPAVQVVERKAT 119
Db 61 ECSQPALMKIKGVNPFVRK-YSDTIAELOLQPSAKDFEVSLVCGHFPAVQVVERKAT 120
Qy 120 GDIYAMKVMKKALLAOEVSFFEEERNILSRSTSPWIPOLQVAFODKNHLYLVEYOPG 179
Db 121 GDIYAMKVMKKALLAOEVSFFEEERNILSRSTSPWIPOLQVAFODKNHLYLVEYOPG 180
Qy 180 GDLSILNRYEDQDENLLOFYLAELITAVSHVLMGVHNDIRENLVTRTGIRKVD 239
Db 181 GDLSILNRYEDQDENLLOFYLAELITAVSHVLMGVHNDIRENLVTRTGIRKVD 240
Qy 240 FGSAAKNSNMVNAKLPITGPDYMAPEVLTVNMGDGKGTGLDGDMMVSGVIAYEMIYG 299
Db 241 FGSAAKNSNMVNAKLPITGPDYMAPEVLTVNMGDGKGTGLDGDMMVSGVIAYEMIYG 300
Qy 300 RSPFAEGTSATFNINMFORFLKFPDDPKVSDFLDIQSLCGQKRLKFEIGCHPF 359
Db 301 RSPFAEGTSATFNINMFORFLKFPDDPKVSDFLDIQSLCGQKRLKFEIGCHPF 360
Qy 360 FSKIDANNIRNSPPPPPTLKSDDTSNFPDEPKNSWSSPQCLSPBEGFGEELPFVGF 419
Db 361 FSKIDANNIRNSPPPPPTLKSDDTSNFPDEPKNSWSSPQCLSPBEGFGEELPFVGF 420
Qy 420 SYSKALGILGRSESVVSGLDSPAKTSSWEKKLITKSKELQDSQDKCHMEDEMTRLHRRV 479
Db 421 SYSKALGILGRSESVVSGLDSPAKTSSWEKKLITKSKELQDSQDKCHMEDEMTRLHRRV 480
Qy 480 SEVEAVLSQKEVELKASGTORSLLEODLATYITECSSLKSLSEQARMEVSQEDDKALQTL 539
Db 481 SEVEAVLSQKEVELKASGTORSLLEODLATYITECSSLKSLSEQARMEVSQEDDKALQTL 540
Qy 540 HDIREQSRKLOEITGEQYQAOVEEMRLMMQOLEBDLYSARRSDLYSESELAESEELAE 599
Db 541 HDIREQSRKLOEITGEQYQAOVEEMRLMMQOLEBDLYSARRSDLYSESELAESEELAE 600
Qy 600 KRKATBCQHKLLKAKDQKPEVGEYALKETNAQOLKIOELQKLEKAVASTAEATL 659
Db 601 KRKATBCQHKLLKAKDQKPEVGEYALKETNAQOLKIOELQKLEKAVASTAEATL 660
Qy 660 QNTRQAKERARELEKLONEDESGIRKKLVEAEERHSLNKKYKLETMERERENRLKD 719
Db 661 QNTRQAKERARELEKLONEDESGIRKKLVEAEERHSLNKKYKLETMERERENRLKD 720
Qy 720 DIOTKSQOIQOMADKILIEBKREAOVSAQHLVHLKQEQHYEEKIKVLINDQIKDOLA 779
Db 721 DIOTKSQOIQOMADKILIEBKREAOVSAQHLVHLKQEQHYEEKIKVLINDQIKDOLA 780

QY 780 DKETLENNMQRHEEBAHEKGIKILSEOKAMINAMDSKIRSLSEORIVELSEANKLAANSLSF 839
 DB 781 DKETLENNMQRHEEBAHEKGIKILSEOKAMINAMDSKIRSLSEORIVELSEANKLAANSLSF 840
 QY 840 TORNNKAOEEMISLSEROCKFYLETOAGLEONKRLKEBOLKISHODSDKORLLEETR 899
 DB 841 TORNNKAOEEMISLSEROCKFYLETOAGLEONKRLKEBOLKISHODSDKORLLEETR 900
 QY 900 LREVSLEHEBOKLEKROLTELQSLQERESQFTALQARAALESQLEQAkteLEETAE 959
 DB 901 LREVSLEHEBOKLEKROLTELQSLQERESQFTALQARAALESQLEQAkteLEETAE 960
 QY 960 AAESEIOALTARDEIORKFDALRNSCTVITDLBOLNQLTEDNALNNQNYLSKQLEDEA 1019
 DB 961 AAESEIOALTARDEIORKFDALRNSCTVITDLBOLNQLTEDNALNNQNYLSKQLEDEA 1020
 QY 1020 SGANDEIVOLRSEVDHLREITEREMQULTSQKQTEALKTCTMLEBQVMDLEALNDEL 1079
 DB 1021 SGANDEIVOLRSEVDHLREITEREMQULTSQKQTEALKTCTMLEBQVMDLEALNDEL 1080
 QY 1080 EKERQWEMRBSVLGDKEKSOFCRVELORMIDTEKQSRARADORTESROVVELAVKEHK 1139
 DB 1081 EKERQWEMRBSVLGDKEKSOFCRVELORMIDTEKQSRARADORTESROVVELAVKEHK 1140
 QY 1140 AEIITALQOALKEOKLKASLSLSDKINDLEKHAMLEMMNARSLOQKLETERELKORLLEBOA 1199
 DB 1141 AEIITALQOALKEOKLKASLSLSDKINDLEKHAMLEMMNARSLOQKLETERELKORLLEBOA 1200
 QY 1200 KLQOQMDLOKXNHIFLITQGLQBALDRADLLKTERSDBLEYOLBNIOVLYSHEKVMKEGITS 1259
 DB 1201 KLQOQMDLOKXNHIFLITQGLQBALDRADLLKTERSDBLEYOLBNIOVLYSHEKVMKEGITS 1260
 QY 1260 QOTKLIDLOKXNMPAKKKKXVPIQYNELKALKEKRCALKEBALKXTEHILRSAREE 1319
 DB 1261 QOTKLIDLOKXNMPAKKKKXVPIQYNELKALKEKRCALKEBALKXTEHILRSAREE 1320
 QY 1320 AAHRKATDHPHSTPATARQOIJAMSAIVRSPEHOPSASMLAPSSRRKESSTPEEFRR 1379
 DB 1321 AAHRKATDHPHSTPATARQOIJAMSAIVRSPEHOPSASMLAPSSRRKESSTPEEFRR 1380
 QY 1380 LKERMHNIIPHRFNVLNMRATKCAVCLDTVHFGQASKLECYQWCHPKCSTCLPATCG 1439
 DB 1381 LKERMHNIIPHRFNVLNMRATKCAVCLDTVHFGQASKLECYQWCHPKCSTCLPATCG 1440
 QY 1440 LPARATHTFAFCGDKNNSPGLQTKERSSSLHLEGMKXVPRNNRGGQGMDRKXIVLEG 1499
 DB 1441 LPARATHTFAFCGDKNNSPGLQTKERSSSLHLEGMKXVPRNNRGGQGMDRKXIVLEG 1500
 QY 1500 SKVLIYDNEABAGORPVEEFELCLPDGVSITHGAVGASELANATAKADVPYILKMHESHPH 1559
 DB 1501 SKVLIYDNEABAGORPVEEFELCLPDGVSITHGAVGASELANATAKADVPYILKMHESHPH 1560
 QY 1560 TTCWFGRTLYLLAPSPDKORWVTALLESVAVAGRVSRKKAADAXLANSLLKLEGGDRL 1619
 DB 1561 TTCWFGRTLYLLAPSPDKORWVTALLESVAVAGRVSRKKAADAXLANSLLKLEGGDRL 1620
 QY 1620 DMNCTLPSSDQVYVGTREGLYALNVLKNSLTHVGTICAVPQIYIIXLEKXLMTAGER 1679
 DB 1621 DMNCTLPSSDQVYVGTREGLYALNVLKNSLTHVGTICAVPQIYIIXLEKXLMTAGER 1680
 QY 1680 ALCLVDVKKVKSQSLAQSHLPAOPDISPNIPEAVKCHLFGAGKINGLCTICAMPSKYVI 1739
 DB 1681 ALCLVDVKKVKSQSLAQSHLPAOPDISPNIPEAVKCHLFGAGKINGLCTICAMPSKYVI 1740
 QY 1740 LRYNENLSKYCIKRIETSEPCSIHFTNYSILIGTNKFEYIDMKQYTLLEFLDKNDHSL 1799
 DB 1741 LRYNENLSKYCIKRIETSEPCSIHFTNYSILIGTNKFEYIDMKQYTLLEFLDKNDHSL 1800
 QY 1800 APAVPAASNSFPVSIVOVNSAGOREBYILCFHEGCVFVDSIGRSRTPDDLKMSLPLAF 1859
 DB 1801 APAVPAASNSFPVSIVOVNSAGOREBYILCFHEGCVFVDSIGRSRTPDDLKMSLPLAF 1860

QY 1860 AYREPYLFTHFNSLEVEIEIOASSAGTPARAYLIDNPPIYGLPAGISSGAILYASSYODK 1919
 DB 1861 AYREPYLFTHFNSLEVEIEIOASSAGTPARAYLIDNPPIYGLPAGISSGAILYASSYODK 1920
 QY 1920 LRVICCKGNLVKESGTEHHRGPPSTRSSPNKRGPTYNHEITKRVASSPAPEGSPHRE 1979
 DB 1921 LRVICCKGNLVKESGTEHHRGPPSTRSSPNKRGPTYNHEITKRVASSPAPEGSPHRE 1980
 QY 1980 PSTHRYREGRTTELIRDKSPGRPIERKSPGRITSTRERSPARLPEDSSRGRLPAGAVR 2039
 DB 1981 PSTHRYREGRTTELIRDKSPGRPIERKSPGRITSTRERSPARLPEDSSRGRLPAGAVR 2040
 QY 2040 TPPLSQVNRKWDQS 2052
 DB 2041 TPPLSQVNRKWDQS 2053
 RESULT 8
 ABG78362
 ID ABG78362 standard; protein; 2066 AA.
 XX ABG78362;
 AC
 XX
 D7 15-NOV-2002 (first entry)
 XX
 DE Human protein, homologous to kinases, designated NOV3A.
 XX
 XX Human; NOV; NOVX-associated disorder; cardiomyopathy;
 KW atherosclerosis; diabetes; cell signaling; metabolic pathway;
 KW cellular receptor; downstream effector; cancer; gene therapy;
 KW hypertension; congenital heart defect; aortic stenosis; obesity;
 KW infectious disease; anorexia; Alzheimer's disease; Parkinson's disease;
 KW neurodegenerative disorder; haemophilia; dyslipidemia; vaccine;
 KW haemopoietic disease; scleroderma; fertility; immunogen;
 KW idiopathic thrombocytopenic purpura; graft versus host disease;
 KW Crohn's disease; multiple sclerosis; cirrhosis; autoimmune disease;
 KW systemic lupus erythematosus; asthma; arthritis; psoriasis; allergy;
 KW stroke; anxiety; Leech-Nyman syndrome; schizophrenia; cerebellar ataxia;
 KW pain; alcoholism; transgenic.
 OS Homo sapiens.
 XX
 XX WO200226826-A2.
 PN
 XX
 PD 04-APR-2002.
 XX
 PE 27-SEP-2001; 2001WO-US042336.
 PR 27-SEP-2000; 2000US-0235631P.
 PR 27-SEP-2000; 2000US-0235631P.
 PR 27-SEP-2000; 2000US-0235808P.
 PR 27-SEP-2000; 2000US-0236064P.
 PR 27-SEP-2000; 2000US-0236065P.
 PR 27-SEP-2000; 2000US-0236066P.
 PR 28-SEP-2000; 2000US-0236135P.
 PR 03-OCT-2000; 2000US-0237434P.
 PR 05-OCT-2000; 2000US-0238321P.
 PR 06-OCT-2000; 2000US-0238399P.
 PR 06-OCT-2000; 2000US-0238399P.
 PR 16-MAR-2001; 2001US-0276667P.
 PR 31-MAY-2001; 2001US-0294823P.
 PR 12-JUL-2001; 2001US-0304868P.
 PR 26-SEP-2001; 2001US-020235631.
 PA (CURA-) CURAGEN CORP.
 XX
 XX Gerlach VL, MacDougall JR, Smithson G, Millet I, Stone D,
 PI Gunther E, Ellerman K, Grosse WM, Alsbrook JP, Lepley DM,
 PI Burgess CE, Padigaru M, Kikuda R, Spytek KA, Leach MD, Shimkets RA;
 DR WPI; 2002-499860/53.
 DR N-Psdb; ABS63435.
 XX

PT Novel isolated NOVX polypeptides and polynucleotides homologous to
PT attractin, plexin, pappin-like family of proteins, useful for treating
PT atherosclerosis, diabetes, cancer, Alzheimer's disease, hemophilia and
PT stroke.

XX Claim 1; Page 42; 308pp; English.

CC The invention discloses the isolated human polypeptides, and
CC polynucleotides encoding them, that have been designated NOVX. The
CC polypeptides, polynucleotides and antibodies are useful in treating or
CC preventing a NOVX-associated disorder which is cardiomyopathy,
CC atherosclerosis and diabetes in a human, where the disorder is related to
CC cell signal processing and metabolic pathway modulation. They can also be
CC used in determining the presence of, or predisposition to, a disease
CC associated with altered levels of the polypeptides and polynucleotides of
CC any one of the 13 sequences (NOV1-NOV13), for raising antibodies, for
CC identifying an agent that binds to, or that modulates the expression or
CC activity of the polypeptide, for identifying an agent which is cellular
CC receptor or downstream effector, for treating or preventing a NOVX-
CC associated disorder and as a pharmaceutical composition comprising the
CC polypeptide, polynucleotide or the antibody. The polypeptides and
CC polynucleotides are useful in diagnostic applications (e.g. as a marker
CC for cancerous cells or tissue types) where their amounts are assessed, or
CC for the manufacture of a medicament (e.g. gene therapy) for treating or
CC preventing disorders or syndromes such as hypertension, congenital heart
CC defects, aortic stenosis, obesity, infectious disease, anorexia, cancer,
CC Alzheimer's disease, Parkinson's disorders, neurodegenerative disorders,
CC hemophilia, idiopathic thrombocytopenic purpura, graft versus host
CC disease, Crohn's disease, multiple sclerosis, cirrhosis, autoimmune
CC disease, systemic lupus erythematosus, asthma, arthritis, psoriasis,
CC allergy, stroke, anxiety, Lesch-Nyhan syndrome, schizophrenia, cerebellar
CC ataxia, pain and alcoholism. They may also be used as immunogens to
CC produce antibodies specific for the invention, and as vaccines.
CC Transgenic cells containing a NOVX expressing construct are useful to
CC produce non-human transgenic animals for studying the function and/or
CC activity of the NOVX proteins and for identifying and/or evaluating
CC modulators of NOVX protein activity. Transgenic cells containing a NOVX
CC expressing construct are useful to produce non-human transgenic animals
CC for studying the function and/or activity of the NOVX proteins and for
CC identifying and/or evaluating modulators of NOVX protein activity. The
CC sequences presented in ABG78359-ABG78371 are the human NOV1-NOV13 proteins
XX
XX
SQ Sequence 2066 AA;

Query Match 99.4%; Score 10425.5; DB 5; Length 2066;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 2044; Conservative 2; Mismatches 5; Indels 1; Gaps 1;

QY 1 MLKRYGARNPLDGAAPISASRLNLPFGKRPFTQOQMSPLSRGILDLFVLF 60
DB 1 MLKRYGARNPLDGAAPISASRLNLPFGKRPFTQOQMSPLSRGILDLFVLF 60
QY 61 ECGOPALMKIKHVNFPKRSYDTTAELOELPSAKDFEVRSLVCGGFAEYQVVRKATG 120
DB 61 ECGOPALMKIKHVNFPKRSYDTTAELOELPSAKDFEVRSLVCGGFAEYQVVRKATG 120
QY 121 DIYAMKVMKKKALLAEOVSFFEEERNILSRSTSPMI POLQYARODKNNHLLVMEYORGG 180
DB 121 DIYAMKVMKKKALLAEOVSFFEEERNILSRSTSPMI POLQYARODKNNHLLVMEYORGG 180
QY 181 DLISLNRVEDOLDENL IQFYLAELILVHSHVLMGVVHRDIKENTILVDRTHGILKLVDF 240
DB 181 DLISLNRVEDOLDENL IQFYLAELILVHSHVLMGVVHRDIKENTILVDRTHGILKLVDF 240
QY 241 GSAAKMNSNMVNAKLIGTPDYNAPEYLTVMGDGGKTYGLDCDWSVGIATVEMTYGR 300
DB 241 GSAAKMNSNMVNAKLIGTPDYNAPEYLTVMGDGGKTYGLDCDWSVGIATVEMTYGR 300
QY 301 SPFAEGTSARTFNINMFORFLKPPDPKVSDFLIDISLLCGCKELKREGLCCHPFF 360
DB 301 SPFAEGTSARTFNINMFORFLKPPDPKVSDFLIDISLLCGCKELKREGLCCHPFF 360

QY 361 SKIDMNNIRNSPPFPVPTLKSDDDTSNPDEPKNSWSSSPQQLSPSGSGEELPFVGRS 420
DB 360 SKIDMNNIRNSPPFPVPTLKSDDDTSNPDEPKNSWSSSPQQLSPSGSGEELPFVGRS 419
QY 421 YSKALGILGRSESVSGSDSPAKTSMEKLLIKSKELQSDQCKHKEOEWTRLHRRVS 480
DB 420 YSKALGILGRSESVSGSDSPAKTSMEKLLIKSKELQSDQCKHKEOEWTRLHRRVS 479
QY 481 EVEAVLSQKEVELKASSETORSLLBDLATYITTECSLSRLSQARMEVSQEDDKALQLIH 540
DB 480 EVEAVLSQKEVELKASSETORSLLBDLATYITTECSLSRLSQARMEVSQEDDKALQLIH 539
QY 541 DIRQSRLQSIKOEYQAOUEEMRLMNOUEEDLVASRRSDLYESILRSRLAAEFK 600
DB 540 DIRQSRLQSIKOEYQAOUEEMRLMNOUEEDLVASRRSDLYESILRSRLAAEFK 599
QY 601 RKATCOGKLLKADQGEVEGEYAKLEKINAEOOLKIOEOLKLEKIVKASTATELLQ 660
DB 600 RKATCOGKLLKADQGEVEGEYAKLEKINAEOOLKIOEOLKLEKIVKASTATELLQ 659
QY 661 NIRAQKRAERELKQNRSDSSEGIKQVLAERHSHLENKYRLTETMERENRLKOD 720
DB 660 NIRAQKRAERELKQNRSDSSEGIKQVLAERHSHLENKYRLTETMERENRLKOD 719
QY 721 IQTSQOIQOQADKILIEEKHRAQVSAQHLVHLKQKEQHYEKKIVLNDQIKQDLAD 780
DB 720 IQTSQOIQOQADKILIEEKHRAQVSAQHLVHLKQKEQHYEKKIVLNDQIKQDLAD 779
QY 781 KETLENNMORHEEAHEKGLTSBOKMINMDSKINSLEBRIYELSRANKLAANSLFT 840
DB 780 KETLENNMORHEEAHEKGLTSBOKMINMDSKINSLEBRIYELSRANKLAANSLFT 839
QY 841 QRNKAQEMISLIRQOKFYLETQAKLEAQRKLEBOLKISIQHSDSKNRLLLELETRL 900
DB 840 QRNKAQEMISLIRQOKFYLETQAKLEAQRKLEBOLKISIQHSDSKNRLLLELETRL 899
QY 901 REVSLHEBQCLEKQUTELQSLQERESQTLQARALBESQRAKTELETTTAA 960
DB 900 REVSLHEBQCLEKQUTELQSLQERESQTLQARALBESQRAKTELETTTAA 959
QY 961 EEEIQTALHARDELQKFPDLRNSCTVITDLEBQNLQTEENALNNQNFSLKQDLAS 1020
DB 960 EEEIQTALHARDELQKFPDLRNSCTVITDLEBQNLQTEENALNNQNFSLKQDLAS 1019
QY 1021 GANDEIVQNRSEVHLRREITEREMQTSQKQTEALKTCTMLEBQVMDLEALNDELLE 1080
DB 1020 GANDEIVQNRSEVHLRREITEREMQTSQKQTEALKTCTMLEBQVMDLEALNDELLE 1079
QY 1081 KERQWEAMRSVLDGKESQFEGRVLEQRMIDTERKOSPARADQRTTESRQVELAVKEHA 1140
DB 1080 KERQWEAMRSVLDGKESQFEGRVLEQRMIDTERKOSPARADQRTTESRQVELAVKEHA 1139
QY 1141 EITLALQALKEQKAKASLSKNDLEKHAMLEMANRSLQOKLETRELEIKORLLEBAK 1200
DB 1140 EITLALQALKEQKAKASLSKNDLEKHAMLEMANRSLQOKLETRELEIKORLLEBAK 1199
QY 1201 LQOQMDLOKNIHIFLTQGLQALRADLLKTERSDLEQYOLENIQVLSHEKVKKEGTSIQ 1260
DB 1200 LQOQMDLOKNIHIFLTQGLQALRADLLKTERSDLEQYOLENIQVLSHEKVKKEGTSIQ 1259
QY 1261 QTKLIDFLQAKMDQPAKKKRVPLQYNEIKLALKEKARCAELBEALQKTRILERSABEEA 1320
DB 1260 QTKLIDFLQAKMDQPAKKKRVPLQYNEIKLALKEKARCAELBEALQKTRILERSABEEA 1319
QY 1321 AHRKATHPHPSTPATARQOQIAMSAYRSPHQSSAMSLAPSSRRKESSTPEEFSRL 1380
DB 1320 AHRKATHPHPSTPATARQOQIAMSAYRSPHQSSAMSLAPSSRRKESSTPEEFSRL 1379
QY 1381 KERQHNNI PHRFNVLGMRATKCAVCLDTYVFGQASKCLEQVMCHPKCSTCLPATGCL 1440
DB 1380 KERQHNNI PHRFNVLGMRATKCAVCLDTYVFGQASKCLEQVMCHPKCSTCLPATGCL 1439
QY 1441 PAEYATHTPEAFCDKXNSPGLQTKPESSLIHLGEMNKVPRNNKRGQOGMDRKYIVLEGS 1500

```

Db      1440 PABYATHTFAFCDDKNMSPGLQTKPESSLHLGMMKVPNNRRGQGGMRKXIVLEGS 1499
Qy      1501 KVLIVDNEARAGORPVBEFELCLPDGVSTHGA VGASELANTKADVPYLLKMSHPHT 1560
Db      1500 KVLIVDNEARAGORPVBEFELCLPDGVSTHGA VGASELANTKADVPYLLKMSHPHT 1559
Qy      1561 TCMFGRLLYLPAFPDKQKRWTALESYVAGRVSRKAEADAKLLGSLKLEGGDRLD 1620
Db      1560 TCMFGRLLYLPAFPDKQKRWTALESYVAGRVSRKAEADAKLLGSLKLEGGDRLD 1619
Qy      1621 MNCITLPPSDQVVLVGTBERGLVAVLVKNSLTNVPICGAVPQIYIIKLEKLMAGERRA 1680
Db      1620 MNCITLPPSDQVVLVGTBERGLVAVLVKNSLTNVPICGAVPQIYIIKLEKLMAGERRA 1679
Qy      1681 LCLVDVKKVKQSLAOSHLPAQPDISPNI FEAVKGCFLFGAKIENGCLICAMPSKVYIL 1740
Db      1680 LCLVDVKKVKQSLAOSHLPAQPDISPNI FEAVKGCFLFGAKIENGCLICAMPSKVYIL 1739
Qy      1741 RYNNELSKYCIKKEIETSEPCSHFTNYSILIGTNKYEIDMKQYTLLEFLLDKNDHSLA 1800
Db      1740 RYNNELSKYCIKKEIETSEPCSHFTNYSILIGTNKYEIDMKQYTLLEFLLDKNDHSLA 1799
Qy      1801 PAVPAASNSFPVSIYOVNSAGORBEYILCFHEGCVFVDSYGRSRDIDLKMSRLPLAFA 1860
Db      1800 PAVPAASNSFPVSIYOVNSAGORBEYILCFHEGCVFVDSYGRSRDIDLKMSRLPLAFA 1859
Qy      1861 YREBYLFTVTHNSLEVEIEIQRSSAGTPARAYLIDPNRYLGPAPASSGATYLASSYODKL 1920
Db      1860 YREBYLFTVTHNSLEVEIEIQRSSAGTPARAYLIDPNRYLGPAPASSGATYLASSYODKL 1919
Qy      1921 RVLCCKGNLVKESGTEHHRGPSSTRSSPNKRGPPTYNEHITKRVASSPAPPEGSHPREP 1980
Db      1920 RVLCCKGNLVKESGTEHHRGPSSTRSSPNKRGPPTYNEHITKRVASSPAPPEGSHPREP 1979
Qy      1981 STPHRYBGRTELARDKSPGRLPRKSPGRILSTRRRSPARLPEDSSRRLPAGAVRT 2040
Db      1980 STPHRYBGRTELARDKSPGRLPRKSPGRILSTRRRSPARLPEDSSRRLPAGAVRT 2039
Qy      2041 PLGQVNVKWDOS 2052
Db      2040 PLGQVNVKWDOS 2051

RESULT 9
ADA05654
ID      ADA05654 standard; procein: 2066 AA.
XX
AC      ADA05654;
XX
DT      06-NOV-2003 (first entry)
XX
DE      Human NOV1g protein SEQ ID NO:14.
XX
KW      human; NOVX; antidiabetic; anorectic; antibacterial; virucide;
KW      immunomodulator; cytosarctic; neuroprotective;
KW      antiparkinsonian; antilipemic; gene therapy; human disease;
KW      metabolic disorder; diabetes; obesity; infection; cachexia; cancer;
KW      neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;
KW      immune disorder; haematopoietic disorder; dyslipidaemia.
XX
OS      Homo sapiens.
XX
PN      M02003029424-A2.
XX
PD      10-APR-2003.
XX
PE      02-OCT-2002; 2002WO-US031373.
XX
PR      02-OCT-2001; 2001US-0326489P.
PR      05-OCT-2001; 2001US-0327435P.
PR      05-OCT-2001; 2001US-0327449P.
PR      09-OCT-2001; 2001US-0327917P.

```

```

PR      09-OCT-2001; 2001US-0328029P.
PR      09-OCT-2001; 2001US-0328044P.
PR      09-OCT-2001; 2001US-0328056P.
PR      12-OCT-2001; 2001US-0328849P.
PR      15-OCT-2001; 2001US-0329414P.
PR      17-OCT-2001; 2001US-0330142P.
PR      18-OCT-2001; 2001US-0330309P.
PR      22-OCT-2001; 2001US-0341058P.
PR      24-OCT-2001; 2001US-0339266P.
PR      24-OCT-2001; 2001US-0343629P.
PR      29-OCT-2001; 2001US-0349575P.
PR      01-NOV-2001; 2001US-0346357P.
PR      17-APR-2002; 2002US-0373260P.
PR      19-APR-2002; 2002US-0373815P.
PR      19-APR-2002; 2002US-0373817P.
PR      19-APR-2002; 2002US-0373826P.
PR      19-APR-2002; 2002US-0373884P.
PR      22-APR-2002; 2002US-0374977P.
PR      16-MAY-2002; 2002US-0381037P.
PR      16-MAY-2002; 2002US-0381038P.
PR      16-MAY-2002; 2002US-0381042P.
PR      17-MAY-2002; 2002US-0381642P.
PR      28-MAY-2002; 2002US-0383656P.
PR      29-MAY-2002; 2002US-0383831P.
PR      25-JUN-2002; 2002US-0391335P.
PR      01-OCT-2002; 2002US-00262511.

XX
XX
PA      (CURA-) CURAGEN CORP.
XX
PI      Smithson G, Millet I, Peyman JA, Kekuda R, Ju J, Li L, Guo X;
PI      Patunajan M, Spytek KA, Edinger SR, Ellerman K, Malyankar UM;
PI      Ort T, Gorman L, Zetunusen BD, Anderson DW, Zhong M, Catterton E;
PI      Ji W, Miller CE, Rastelli L, Stone DJ, Pena CE, Shenoy SG;
PI      Shmeks RA, Rothenberg ME, Leach MD, Agee ML, Berghs C, DiIippo VA;
PI      Eisen AJ, Gangolli BA, Rieger DK, Spaderna SK;
XX
DR      WPI; 2003-381626/36.
XX
XX      N-PSDB; ADA05653.
XX
PT      New NOVX polypeptides and nucleic acids, useful for diagnosing,
PT      preventing or treating NOVX-associated disorders, e.g. diabetes, obesity,
PT      cancer or dyslipidemia, and in chromosome mapping, tissue typing or
PT      pharmacogenomics.
XX
XX
PS      Claim 1; Page 105-106; 586pp; English.
XX
XX
XX      The present invention describes NOVX proteins, where X can be 1 to 55
XX      (e.g. NOV1). Also described: (1) a composition comprising a polypeptide
XX      described above and a carrier; (2) a kit comprising, in one or more
XX      containers, the composition described above; (3) an isolated nucleic acid
XX      molecule which encodes a NOVX protein of the invention; (4) a vector
XX      comprising the nucleic acid molecule described above; (5) a cell
XX      comprising the above vector; (6) an antibody that immunospecifically
XX      binds to the polypeptide described above; (7) methods for determining the
XX      presence or amount of the above polypeptide or nucleic acid molecule in a
XX      sample; (8) methods for determining the presence of or predisposition to
XX      a disease associated with altered levels of expression of the above
XX      polypeptide or nucleic acid molecule in a first mammalian subject; (9) a
XX      method of identifying an agent that binds to the polypeptide described
XX      above; (10) a method for identifying a potential therapeutic agent for
XX      use in treating a pathology that is related to an aberrant expression or
XX      aberrant physiological interactions of the polypeptide; (11) a method of
XX      screening for a modulator of activity or of latency or predisposition to
XX      a pathology associated with the polypeptide; (12) a method for modulating
XX      the activity of the polypeptide described above; (13) methods of treating
XX      or preventing a pathology associated with the above polypeptide in a
XX      mammal; and (14) a method for producing the above polypeptide. NOVX
XX      sequences have antidiabetic, anorectic, antibacterial, virucide,
XX      immunomodulator, cytosarctic, neuroprotective, antiparkinsonian
XX      and antilipemic activities, and can be used in gene therapy. The
XX      polypeptide is useful in manufacturing a medicament for treating a
XX      syndrome associated with a human disease. The polypeptide or the nucleic
XX      acid molecule may be used to diagnose, treat or prevent metabolic

```

CC disorders such as diabetes or obesity, infections, cachexia, cancer,
CC neurodegenerative disorders such as Alzheimer's disease or Parkinson's
CC disease, immune disorders, haematopoietic disorders and various
CC dyplidemia. The nucleic acids can also be used as hybridisation
CC probes, in chromosome mapping, tissue typing, preventive medicine and
CC pharmacogenomics. The present sequence represents a human NOVX from the
CC present invention.

XX Sequence 2066 AA;

Query Match 99.4%; Score 10425.5; DB 6; Length 2066;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 2044; Conservative 2; Mismatches 5; Indels 1; Gaps 1;

Qy 1 MLKFKYGRNPLDGAAPISASRLNLPFGKPPFTQOMSPISREGILDLFVLE 60
Db 1 MLKFKYGRNPLDGAAPISASRLNLPFGKPPFTQOMSPISREGILDLFVLE 60
Qy 61 ECSOPALMKIKNVSNFPAKISDTTAELOEPSAKDFVRSLVGGHFAEVQVYREKATG 120
Db 61 ECSOPALMKIKNVSNFPAKISDTTAELOEPSAKDFVRSLVGGHFAEVQVYREKATG 120
Qy 121 DIYAMKWKKKKALLAEOVSFFEEERNLISRTSPMTIQLQYAFODKXHLVLMEXOPGG 180
Db 121 DIYAMKWKKKKALLAEOVSFFEEERNLISRTSPMTIQLQYAFODKXHLVLMEXOPGG 180
Qy 181 DLSLNRVEDOLDENLIQFYLAETILAVSHVLMGYVHRDIKPEIILVDRTHIKLVDF 240
Db 181 DLSLNRVEDOLDENLIQFYLAETILAVSHVLMGYVHRDIKPEIILVDRTHIKLVDF 240
Qy 241 GSAAKNSNKNVNAKLPIGTDYAPAEVLTVMGDGGKTYGLDCDWSVGVIAYEMTYGR 300
Db 241 GSAAKNSNKNVNAKLPIGTDYAPAEVLTVMGDGGKTYGLDCDWSVGVIAYEMTYGR 300
Qy 301 SPFAEGTSARFNNINMFORFLKRPDPKVSDFLDLIQSLCGOKERLKEGICCHPF 360
Db 301 SPFAEGTSARFNNINMFORFLKRPDPKVSDFLDLIQSLCGOKERLKEGICCHPF 360
Qy 361 SKIDMNNIRNPPFPVFTLKSDDTSNFDPEKNSWVSSPCQSPSGFSEELPFVGF 420
Db 361 SKIDMNNIRNPPFPVFTLKSDDTSNFDPEKNSWVSSPCQSPSGFSEELPFVGF 420
Qy 421 YSKALGILGRSESVVSGIDSPAKTSMEKLLIKSKELQDSQDCHKMEQDMTRLHRVS 480
Db 421 YSKALGILGRSESVVSGIDSPAKTSMEKLLIKSKELQDSQDCHKMEQDMTRLHRVS 480
Qy 481 EYEVAVLSQKVELQASTONSLEODLATYITTECSLSKSLERQARMEVSOQDYALQLH 540
Db 481 EYEVAVLSQKVELQASTONSLEODLATYITTECSLSKSLERQARMEVSOQDYALQLH 540
Qy 541 DIRQSRKLOEIKEOEYOAVEEMRLMNNOLEEDLVASRRSDIYSESELRESRLAABEF 600
Db 541 DIRQSRKLOEIKEOEYOAVEEMRLMNNOLEEDLVASRRSDIYSESELRESRLAABEF 600
Qy 601 RKATCECHKLKADQKPEVGEYAKLEKINAEOQLKIOLEQLEKRAVKASTEATELLQ 660
Db 601 RKATCECHKLKADQKPEVGEYAKLEKINAEOQLKIOLEQLEKRAVKASTEATELLQ 660
Qy 661 NIROAKRARELEKLNREDSSEIGIRKCLVEABERHSLNENKYLTERRENNRLKDD 720
Db 661 NIROAKRARELEKLNREDSSEIGIRKCLVEABERHSLNENKYLTERRENNRLKDD 720
Qy 721 IOTSSQOIQOQADKILEEKGREAOVSAOHLAEVHLKOEQHYBEKI KVLINOIKKDLAD 780
Db 721 IOTSSQOIQOQADKILEEKGREAOVSAOHLAEVHLKOEQHYBEKI KVLINOIKKDLAD 780
Qy 781 KETLENNMORHEEBAHEKGLISSEOKAMINAMDSKISLEORIVELSEANGLAANSILFT 840
Db 781 KETLENNMORHEEBAHEKGLISSEOKAMINAMDSKISLEORIVELSEANGLAANSILFT 840
Qy 841 ORNKAQOEMISELROOKFYLETQAGKLEAONRKLBEOLLEKISHQDSHDKNRLLELETRL 900
Db 841 ORNKAQOEMISELROOKFYLETQAGKLEAONRKLBEOLLEKISHQDSHDKNRLLELETRL 900

Qy 901 REVSLEHEEQKLEIKRQLTTELQSLQERESQVTLQAARALBESQLRAKTELEETTAA 960
Db 901 REVSLEHEEQKLEIKRQLTTELQSLQERESQVTLQAARALBESQLRAKTELEETTAA 960
Qy 961 EEEIOALTARHDEIQRKFDALRNSCTVITDLEQONCUTENNAELNNOFPYLSKQLDAS 1020
Db 961 EEEIOALTARHDEIQRKFDALRNSCTVITDLEQONCUTENNAELNNOFPYLSKQLDAS 1020
Qy 1021 GANDEIVOLRSEVHLRREITEREMQTSQOKTMEALKTTCQTMLEBOVMDEALNDELLE 1080
Db 1021 GANDEIVOLRSEVHLRREITEREMQTSQOKTMEALKTTCQTMLEBOVMDEALNDELLE 1080
Qy 1081 KERQWEAMRSVLGDEKQFECRVALEQLRMLDTEKQSRARADQRTTESROVELAVKEKA 1140
Db 1081 KERQWEAMRSVLGDEKQFECRVALEQLRMLDTEKQSRARADQRTTESROVELAVKEKA 1140
Qy 1141 EILALQALKEQKIKASLSLKLNDLEKKAMLENNASLQOKLETETELKORLLEBOAK 1200
Db 1141 EILALQALKEQKIKASLSLKLNDLEKKAMLENNASLQOKLETETELKORLLEBOAK 1200
Qy 1200 LQOQMDLOKNIIFRLTQGLQALDRADLKTERSDLEYQLENIQVLYSHEKYKMEGTISQ 1260
Db 1200 LQOQMDLOKNIIFRLTQGLQALDRADLKTERSDLEYQLENIQVLYSHEKYKMEGTISQ 1260
Qy 1261 QTKLIDFLQAMMDPAKKKKVPLQYNELKALKEKARCALBEALQKTRIELSAREEA 1320
Db 1261 QTKLIDFLQAMMDPAKKKKVPLQYNELKALKEKARCALBEALQKTRIELSAREEA 1320
Qy 1321 AHRKATHPHPSTPATARQOITAMSAIVRSPEHOPSMSLAPSSRRKESSTPEEFSSRL 1380
Db 1321 AHRKATHPHPSTPATARQOITAMSAIVRSPEHOPSMSLAPSSRRKESSTPEEFSSRL 1380
Qy 1381 KERHNNHNPFRFNGLMARATKCAVCLDTVAFGQASCLCEQWCHPKCSTCLPATGL 1440
Db 1381 KERHNNHNPFRFNGLMARATKCAVCLDTVAFGQASCLCEQWCHPKCSTCLPATGL 1440
Qy 1441 PAEYATFTTEAFCDKKNNSPGLQTKEPSSSLHLEGMKVPNNRRGOQGMKRYIVLEGS 1500
Db 1441 PAEYATFTTEAFCDKKNNSPGLQTKEPSSSLHLEGMKVPNNRRGOQGMKRYIVLEGS 1500
Qy 1500 KVLIIYDNEAREAGRPVEBEFLCLPDGDVSIHGAVGASELANAKADVPIYLKMSHPHT 1560
Db 1500 KVLIIYDNEAREAGRPVEBEFLCLPDGDVSIHGAVGASELANAKADVPIYLKMSHPHT 1560
Qy 1561 TCWPGRTLYLLAPSPDKORVNTALBESVAGRSREKAEADAKLGNLSLKLREGDDRLD 1620
Db 1561 TCWPGRTLYLLAPSPDKORVNTALBESVAGRSREKAEADAKLGNLSLKLREGDDRLD 1620
Qy 1621 MNCTLPPSDQVVLVGTSEGLYALNVLNKSLTRHVGIGAVFOIYIILKOLEKLMTAGERA 1680
Db 1621 MNCTLPPSDQVVLVGTSEGLYALNVLNKSLTRHVGIGAVFOIYIILKOLEKLMTAGERA 1680
Qy 1680 LCLVDVKKVQSLAQSHLPAPDISPNIPEAVKCHLFGAKIENGCLICAMPSSKVYL 1740
Db 1680 LCLVDVKKVQSLAQSHLPAPDISPNIPEAVKCHLFGAKIENGCLICAMPSSKVYL 1740
Qy 1741 RYNNENLKYCIKREIETSEPCSTHFTNYSILITGNKFYEIDMKOYTLLEFLDKNDHSLA 1800
Db 1741 RYNNENLKYCIKREIETSEPCSTHFTNYSILITGNKFYEIDMKOYTLLEFLDKNDHSLA 1800
Qy 1801 PAVPAASNSPPVSIYOVNSAGREBYLLCHHEGVFVDSYGRSRDIDLKMSRLPLAFA 1860
Db 1801 PAVPAASNSPPVSIYOVNSAGREBYLLCHHEGVFVDSYGRSRDIDLKMSRLPLAFA 1860
Qy 1860 YREBYLFTVTHNSLEVIEIQARSAGTPAPAYLIDIPRVLGPAISSGAIYLAASYQKL 1920
Db 1860 YREBYLFTVTHNSLEVIEIQARSAGTPAPAYLIDIPRVLGPAISSGAIYLAASYQKL 1920
Qy 1920 RVICCKNLVKESTGTHRRGPSTSSPNKRGPTVNEHTTKRAVASSPAPPEGSHPREP 1979
Db 1920 RVICCKNLVKESTGTHRRGPSTSSPNKRGPTVNEHTTKRAVASSPAPPEGSHPREP 1979

Matches 2043; Conservative 2; Mismatches 8; Indels 1; Gaps 1;

QY 1 MLKRTGARNPLDGAAPLISRASRLNLFQOKPPEMTQOQMSPLSEGLIDALFVLE 60
Db 1 MLKRTGARNPLDGAAPLISRASRLNLFQOKPPEMTQOQMSPLSEGLIDALFVLE 60
QY 61 BCSOPALMKIKHVNFPVKYSDTIAELOPSAKDFVRSVINGCGHAEQVYREKATG 120
Db 61 BCSOPALMKIKHVNFPVKYSDTIAELOPSAKDFVRSVINGCGHAEQVYREKATG 120
QY 61 BCSOPALMKIKHVNFPVKYSDTIAELOPSAKDFVRSVINGCGHAEQVYREKATG 120
Db 61 BCSOPALMKIKHVNFPVKYSDTIAELOPSAKDFVRSVINGCGHAEQVYREKATG 120
QY 121 DIYAMKVMKKKALLAOEVSFFEEERNILSRSTPMIQLQYAFQDKHLLVMEYORG 180
Db 121 DIYAMKVMKKKALLAOEVSFFEEERNILSRSTPMIQLQYAFQDKHLLVMEYORG 180
QY 121 DIYAMKVMKKKALLAOEVSFFEEERNILSRSTPMIQLQYAFQDKHLLVMEYORG 180
Db 121 DIYAMKVMKKKALLAOEVSFFEEERNILSRSTPMIQLQYAFQDKHLLVMEYORG 180
QY 181 DLSLNRYEDQDLENTIQFYLAELILAVSHVLMGVYHARDIKENILVDRTHIKLVDF 240
Db 181 DLSLNRYEDQDLENTIQFYLAELILAVSHVLMGVYHARDIKENILVDRTHIKLVDF 240
QY 241 GSAAKMNSNMVNAKPLIGTPDYAPVYLTMNGDGKTYGLDCDMSVGVYIAYEMTYGR 300
Db 241 GSAAKMNSNMVNAKPLIGTPDYAPVYLTMNGDGKTYGLDCDMSVGVYIAYEMTYGR 300
QY 241 GSAAKMNSNMVNAKPLIGTPDYAPVYLTMNGDGKTYGLDCDMSVGVYIAYEMTYGR 300
Db 241 GSAAKMNSNMVNAKPLIGTPDYAPVYLTMNGDGKTYGLDCDMSVGVYIAYEMTYGR 300
QY 301 SPFAEGSARTFNINMFORFLKFPDDPKVSDFLDIQSLGCKEKLKEEGCCHPEF 360
Db 301 SPFAEGSARTFNINMFORFLKFPDDPKVSDFLDIQSLGCKEKLKEEGCCHPEF 360
QY 301 SPFAEGSARTFNINMFORFLKFPDDPKVSDFLDIQSLGCKEKLKEEGCCHPEF 360
Db 301 SPFAEGSARTFNINMFORFLKFPDDPKVSDFLDIQSLGCKEKLKEEGCCHPEF 360
QY 361 SKIDMNNIRNSPPFVPTLSDDDTSNFDPEKNSWVSSPCQSPGSGSEELPFVQFS 420
Db 361 SKIDMNNIRNSPPFVPTLSDDDTSNFDPEKNSWVSSPCQSPGSGSEELPFVQFS 420
QY 361 SKIDMNNIRNSPPFVPTLSDDDTSNFDPEKNSWVSSPCQSPGSGSEELPFVQFS 420
Db 361 SKIDMNNIRNSPPFVPTLSDDDTSNFDPEKNSWVSSPCQSPGSGSEELPFVQFS 420
QY 421 YSKALGILGNSESVSGSDSPAKTSSWEKLLIKSKELQSDQCHKMOEMTRILHRVS 480
Db 421 YSKALGILGNSESVSGSDSPAKTSSWEKLLIKSKELQSDQCHKMOEMTRILHRVS 480
QY 421 YSKALGILGNSESVSGSDSPAKTSSWEKLLIKSKELQSDQCHKMOEMTRILHRVS 480
Db 421 YSKALGILGNSESVSGSDSPAKTSSWEKLLIKSKELQSDQCHKMOEMTRILHRVS 480
QY 481 EWEAVLSOKEVELKASETORSILFODLATYTCSSSLKSLSEAKRMVSOQDDALQULH 540
Db 481 EWEAVLSOKEVELKASETORSILFODLATYTCSSSLKSLSEAKRMVSOQDDALQULH 540
QY 481 EWEAVLSOKEVELKASETORSILFODLATYTCSSSLKSLSEAKRMVSOQDDALQULH 540
Db 481 EWEAVLSOKEVELKASETORSILFODLATYTCSSSLKSLSEAKRMVSOQDDALQULH 540
QY 541 DIRBOSKLOEIKOEYOAOVEEMRLMNOLEEDLVASRRSDLYESELRESRLAAEFK 600
Db 541 DIRBOSKLOEIKOEYOAOVEEMRLMNOLEEDLVASRRSDLYESELRESRLAAEFK 600
QY 541 DIRBOSKLOEIKOEYOAOVEEMRLMNOLEEDLVASRRSDLYESELRESRLAAEFK 600
Db 541 DIRBOSKLOEIKOEYOAOVEEMRLMNOLEEDLVASRRSDLYESELRESRLAAEFK 600
QY 601 RKAITECHKLLKAKOQKPEVGEYAKLEKINAEQOLKIQELQELKRAVASTATELQ 660
Db 601 RKAITECHKLLKAKOQKPEVGEYAKLEKINAEQOLKIQELQELKRAVASTATELQ 660
QY 601 RKAITECHKLLKAKOQKPEVGEYAKLEKINAEQOLKIQELQELKRAVASTATELQ 660
Db 601 RKAITECHKLLKAKOQKPEVGEYAKLEKINAEQOLKIQELQELKRAVASTATELQ 660
QY 661 NIQAOKARARELEKONREDSSEGTIRKULVEABERHSHLENKYLPTMERRENRKLD 720
Db 661 NIQAOKARARELEKONREDSSEGTIRKULVEABERHSHLENKYLPTMERRENRKLD 720
QY 661 NIQAOKARARELEKONREDSSEGTIRKULVEABERHSHLENKYLPTMERRENRKLD 720
Db 661 NIQAOKARARELEKONREDSSEGTIRKULVEABERHSHLENKYLPTMERRENRKLD 720
QY 721 IOTKSQOIQOMADKILEEKHREAOVSAOHLVEHLKQEKQHYEKKIKVLDOIKKDLAD 780
Db 721 IOTKSQOIQOMADKILEEKHREAOVSAOHLVEHLKQEKQHYEKKIKVLDOIKKDLAD 780
QY 721 IOTKSQOIQOMADKILEEKHREAOVSAOHLVEHLKQEKQHYEKKIKVLDOIKKDLAD 780
Db 721 IOTKSQOIQOMADKILEEKHREAOVSAOHLVEHLKQEKQHYEKKIKVLDOIKKDLAD 780
QY 781 KETLENNMOGHEEBAHEKGIKILSEQKAMINAMDSKIRSLBORIVELSEANKLAANSLSFT 840
Db 781 KETLENNMOGHEEBAHEKGIKILSEQKAMINAMDSKIRSLBORIVELSEANKLAANSLSFT 840
QY 781 KETLENNMOGHEEBAHEKGIKILSEQKAMINAMDSKIRSLBORIVELSEANKLAANSLSFT 840
Db 781 KETLENNMOGHEEBAHEKGIKILSEQKAMINAMDSKIRSLBORIVELSEANKLAANSLSFT 840
QY 841 QRMWKAQOEMISELROQKFLYETOAGKLEAONKRLSEOLEKISIQHSDSDKRRILLETRL 900
Db 841 QRMWKAQOEMISELROQKFLYETOAGKLEAONKRLSEOLEKISIQHSDSDKRRILLETRL 900
QY 841 QRMWKAQOEMISELROQKFLYETOAGKLEAONKRLSEOLEKISIQHSDSDKRRILLETRL 900
Db 841 QRMWKAQOEMISELROQKFLYETOAGKLEAONKRLSEOLEKISIQHSDSDKRRILLETRL 900
QY 901 REVSLHEEOQKLEKROLTELQSLQERESQLTALQARALAESQLQOAKTELEETTAFA 960
Db 901 REVSLHEEOQKLEKROLTELQSLQERESQLTALQARALAESQLQOAKTELEETTAFA 960
QY 901 REVSLHEEOQKLEKROLTELQSLQERESQLTALQARALAESQLQOAKTELEETTAFA 960
Db 901 REVSLHEEOQKLEKROLTELQSLQERESQLTALQARALAESQLQOAKTELEETTAFA 960
QY 961 EEEIQTATARDEIQRFKDALRNSCTVITDLEEQNLQTEBDNAELNNQNFYLSKQDSEAS 1020
Db 961 EEEIQTATARDEIQRFKDALRNSCTVITDLEEQNLQTEBDNAELNNQNFYLSKQDSEAS 1020
QY 961 EEEIQTATARDEIQRFKDALRNSCTVITDLEEQNLQTEBDNAELNNQNFYLSKQDSEAS 1020
Db 961 EEEIQTATARDEIQRFKDALRNSCTVITDLEEQNLQTEBDNAELNNQNFYLSKQDSEAS 1020
QY 1021 GANDEIYQLASEVDHLAREITEREMOULTSQOTMEALKTCTMLEEQVMDLEALNDELLE 1080
Db 1021 GANDEIYQLASEVDHLAREITEREMOULTSQOTMEALKTCTMLEEQVMDLEALNDELLE 1080
QY 1021 GANDEIYQLASEVDHLAREITEREMOULTSQOTMEALKTCTMLEEQVMDLEALNDELLE 1080
Db 1021 GANDEIYQLASEVDHLAREITEREMOULTSQOTMEALKTCTMLEEQVMDLEALNDELLE 1080

QY 1081 KERQWEAMRSVULGDEKQFECRVERBELQMLDTEKQSRARADQRTSESQVVELAVKEHKA 1140
Db 1081 KERQWEAMRSVULGDEKQFECRVERBELQMLDTEKQSRARADQRTSESQVVELAVKEHKA 1140
QY 1080 KERQWEAMRSVULGDEKQFECRVERBELQMLDTEKQSRARADQRTSESQVVELAVKEHKA 1139
Db 1080 KERQWEAMRSVULGDEKQFECRVERBELQMLDTEKQSRARADQRTSESQVVELAVKEHKA 1139
QY 1141 EILALQOALKQOKIKASISLQKNDLEKHAMLENNASLSQOKLETTERELQRLLEEQAK 1200
Db 1141 EILALQOALKQOKIKASISLQKNDLEKHAMLENNASLSQOKLETTERELQRLLEEQAK 1200
QY 1140 EILALQOALKQOKIKASISLQKNDLEKHAMLENNASLSQOKLETTERELQRLLEEQAK 1199
Db 1140 EILALQOALKQOKIKASISLQKNDLEKHAMLENNASLSQOKLETTERELQRLLEEQAK 1199
QY 1201 LOQOMDLOKNIHIFLUTQLOALRADLLKTERSDLEQLENIQVLYSHEKVKMEGTISQ 1260
Db 1201 LOQOMDLOKNIHIFLUTQLOALRADLLKTERSDLEQLENIQVLYSHEKVKMEGTISQ 1260
QY 1200 LOQOMDLOKNIHIFLUTQLOALRADLLKTERSDLEQLENIQVLYSHEKVKMEGTISQ 1259
Db 1200 LOQOMDLOKNIHIFLUTQLOALRADLLKTERSDLEQLENIQVLYSHEKVKMEGTISQ 1259
QY 1261 QTKLIDFLQAKMDQPAKKKKVPLQYNELKALKEKAKCALEBQALQTRILELSAREBA 1320
Db 1261 QTKLIDFLQAKMDQPAKKKKVPLQYNELKALKEKAKCALEBQALQTRILELSAREBA 1320
QY 1260 QTKLIDFLQAKMDQPAKKKKVPLQYNELKALKEKAKCALEBQALQTRILELSAREBA 1319
Db 1260 QTKLIDFLQAKMDQPAKKKKVPLQYNELKALKEKAKCALEBQALQTRILELSAREBA 1319
QY 1321 AHRKATDHPPHSTATATARQOJLMSALYRSPEHQSAMSLLAPSSRRKESSTPEEFSTR 1380
Db 1321 AHRKATDHPPHSTATATARQOJLMSALYRSPEHQSAMSLLAPSSRRKESSTPEEFSTR 1380
QY 1320 AHRKATDHPPHSTATATARQOJLMSALYRSPEHQSAMSLLAPSSRRKESSTPEEFSTR 1379
Db 1320 AHRKATDHPPHSTATATARQOJLMSALYRSPEHQSAMSLLAPSSRRKESSTPEEFSTR 1379
QY 1381 KERHNNHNI PHRFNVLGNMRAATKCAVCLDTYHFGQASKCEQWCHPKGCTCLPATG 1440
Db 1381 KERHNNHNI PHRFNVLGNMRAATKCAVCLDTYHFGQASKCEQWCHPKGCTCLPATG 1440
QY 1380 KERHNNHNI PHRFNVLGNMRAATKCAVCLDTYHFGQASKCEQWCHPKGCTCLPATG 1439
Db 1380 KERHNNHNI PHRFNVLGNMRAATKCAVCLDTYHFGQASKCEQWCHPKGCTCLPATG 1439
QY 1441 PAEYATHTTEAFCDKKNMSPGLOTKEPSSSLHLEGMKVPANNRQGOQMDRKTYVLEGS 1500
Db 1441 PAEYATHTTEAFCDKKNMSPGLOTKEPSSSLHLEGMKVPANNRQGOQMDRKTYVLEGS 1500
QY 1440 PAEYATHTTEAFCDKKNMSPGLOTKEPSSSLHLEGMKVPANNRQGOQMDRKTYVLEGS 1499
Db 1440 PAEYATHTTEAFCDKKNMSPGLOTKEPSSSLHLEGMKVPANNRQGOQMDRKTYVLEGS 1499
QY 1501 KVLTYNDENBARAGORPYEEFELCLPDGDVSIHGAVGASELANAKADVPYILKMSHPHT 1560
Db 1501 KVLTYNDENBARAGORPYEEFELCLPDGDVSIHGAVGASELANAKADVPYILKMSHPHT 1560
QY 1500 KVLTYNDENBARAGORPYEEFELCLPDGDVSIHGAVGASELANAKADVPYILKMSHPHT 1559
Db 1500 KVLTYNDENBARAGORPYEEFELCLPDGDVSIHGAVGASELANAKADVPYILKMSHPHT 1559
QY 1561 TCMFGRITLYLLAPFPKORVNTALLESVAGGRSREKABADKILGNSLLKSGDDBLD 1620
Db 1561 TCMFGRITLYLLAPFPKORVNTALLESVAGGRSREKABADKILGNSLLKSGDDBLD 1620
QY 1560 TCMFGRITLYLLAPFPKORVNTALLESVAGGRSREKABADKILGNSLLKSGDDBLD 1619
Db 1560 TCMFGRITLYLLAPFPKORVNTALLESVAGGRSREKABADKILGNSLLKSGDDBLD 1619
QY 1621 MNCTLPSPSDQVVLVGTGEGYALANVLKNSLTHVAGVFOIYIILKOLEKLMAGEERA 1680
Db 1621 MNCTLPSPSDQVVLVGTGEGYALANVLKNSLTHVAGVFOIYIILKOLEKLMAGEERA 1680
QY 1620 MNCTLPSPSDQVVLVGTGEGYALANVLKNSLTHVAGVFOIYIILKOLEKLMAGEERA 1679
Db 1620 MNCTLPSPSDQVVLVGTGEGYALANVLKNSLTHVAGVFOIYIILKOLEKLMAGEERA 1679
QY 1681 LCLVDVKKVQOSLAQSHLPAPODISEPNI FEAVKGCCHLFGAGKINGLCTCAMPSKVYL 1740
Db 1681 LCLVDVKKVQOSLAQSHLPAPODISEPNI FEAVKGCCHLFGAGKINGLCTCAMPSKVYL 1740
QY 1680 LCLVDVKKVQOSLAQSHLPAPODISEPNI FEAVKGCCHLFGAGKINGLCTCAMPSKVYL 1739
Db 1680 LCLVDVKKVQOSLAQSHLPAPODISEPNI FEAVKGCCHLFGAGKINGLCTCAMPSKVYL 1739
QY 1741 RYNNENLSKYCIKREIETSEPCSHFTNYSTILGTMKFYEIDMKQYTLLEFLDKNDHSLA 1800
Db 1741 RYNNENLSKYCIKREIETSEPCSHFTNYSTILGTMKFYEIDMKQYTLLEFLDKNDHSLA 1800
QY 1740 RYNNENLSKYCIKREIETSEPCSHFTNYSTILGTMKFYEIDMKQYTLLEFLDKNDHSLA 1799
Db 1740 RYNNENLSKYCIKREIETSEPCSHFTNYSTILGTMKFYEIDMKQYTLLEFLDKNDHSLA 1799
QY 1801 PAVFAASNSFPVSIYOVNSAGOREEYILCFHEFGVFVDSYGRSRTDDLKMSRLPLAFA 1860
Db 1801 PAVFAASNSFPVSIYOVNSAGOREEYILCFHEFGVFVDSYGRSRTDDLKMSRLPLAFA 1860
QY 1800 PAVFAASNSFPVSIYOVNSAGOREEYILCFHEFGVFVDSYGRSRTDDLKMSRLPLAFA 1859
Db 1800 PAVFAASNSFPVSIYOVNSAGOREEYILCFHEFGVFVDSYGRSRTDDLKMSRLPLAFA 1859
QY 1861 YREBYLFVTHFNLSLEVIEIQARSSAGTPARAYLDIPNRYLGPALISSGATYLAASYODKL 1920
Db 1861 YREBYLFVTHFNLSLEVIEIQARSSAGTPARAYLDIPNRYLGPALISSGATYLAASYODKL 1920
QY 1860 YREBYLFVTHFNLSLEVIEIQARSSAGTPARAYLDIPNRYLGPALISSGATYLAASYODKL 1919
Db 1860 YREBYLFVTHFNLSLEVIEIQARSSAGTPARAYLDIPNRYLGPALISSGATYLAASYODKL 1919
QY 1921 RVLCCKGNLYKESGTEHHRGPSTSRSSPNKRGPPTYENHTTKRVASSPAPEGSHRREP 1980
Db 1921 RVLCCKGNLYKESGTEHHRGPSTSRSSPNKRGPPTYENHTTKRVASSPAPEGSHRREP 1980
QY 1920 RVLCCKGNLYKESGTEHHRGPSTSRSSPNKRGPPTYENHTTKRVASSPAPEGSHRREP 1979
Db 1920 RVLCCKGNLYKESGTEHHRGPSTSRSSPNKRGPPTYENHTTKRVASSPAPEGSHRREP 1979
QY 1981 STPHRYEGRTELRRDSSPGRLERKSPGRLSTRERSPARLFEQSSSRGLPAGAVRT 2040
Db 1981 STPHRYEGRTELRRDSSPGRLERKSPGRLSTRERSPARLFEQSSSRGLPAGAVRT 2040
QY 2040 PLSQVNVMDQSSV 2054
Db 2040 PLSQVNVMDQSSV 2053
QY 2041 PLSQVNVMDQSSV 2054
Db 2041 PLSQVNVMDQSSV 2053

RESULT 11
ADN63228
ID ADN63228 standard; protein; 2053 AA.

```

XX AC ADN63228;
XX
XX 01-JUN-2004 (first entry)
XX
XX Human NOVA variant.
XX
XX human; NOVA; metabolic disorder; diabetes; obesity; infectious disease;
XX anorexia; cancer; cancer-associated cachexia; neurodegenerative disorder;
XX Alzheimer's disease; Parkinson's disease; immune disorder;
XX hematoepioretic disorder; dyslipidaemia; metabolic syndrome X;
XX wasting disorder.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX Misc-difference 1440
XX /note= "Pro substituted by Leu as a result of single
XX nucleotide polymorphism"
XX
XX US2004038223-A1.
XX
XX 26-FEB-2004.
XX
XX 01-OCT-2002; 2002US-00262511.
XX
XX 02-OCT-2001; 2001US-0326483P.
XX 05-OCT-2001; 2001US-0327435P.
XX 05-OCT-2001; 2001US-0327449P.
XX 09-OCT-2001; 2001US-0327917P.
XX 09-OCT-2001; 2001US-0328029P.
XX 09-OCT-2001; 2001US-0328049P.
XX 09-OCT-2001; 2001US-0328056P.
XX 12-OCT-2001; 2001US-0328849P.
XX 15-OCT-2001; 2001US-0329414P.
XX 17-OCT-2001; 2001US-0330142P.
XX 18-OCT-2001; 2001US-0330309P.
XX 22-OCT-2001; 2001US-0341058P.
XX 24-OCT-2001; 2001US-0339266P.
XX 24-OCT-2001; 2001US-0343629P.
XX 29-OCT-2001; 2001US-0349575P.
XX 01-NOV-2001; 2001US-0346357P.
XX 17-APR-2002; 2002US-0373260P.
XX 19-APR-2002; 2002US-0373815P.
XX 19-APR-2002; 2002US-0373817P.
XX 19-APR-2002; 2002US-0373826P.
XX 19-APR-2002; 2002US-0373884P.
XX 22-APR-2002; 2002US-0374977P.
XX 16-MAY-2002; 2002US-0381037P.
XX 16-MAY-2002; 2002US-0381038P.
XX 16-MAY-2002; 2002US-0381042P.
XX 17-MAY-2002; 2002US-0381642P.
XX 28-MAY-2002; 2002US-0383656P.
XX 29-MAY-2002; 2002US-0383831P.
XX 25-JUN-2002; 2002US-0391335P.
XX
XX (SMIT/) SMITHSON G.
XX (MILT/) MILLET I.
XX (PEYM/) PEYMAN J A.
XX (KEKU/) KEKUDA R.
XX (JUUJ/) JU J.
XX (LILL/) LI L.
XX (GUOX/) GUO X.
XX (PATU/) PATURAJAN M.
XX (SPYT/) SPYTEK K A.
XX (EDIN/) EDINGER S R.
XX (ELLE/) ELLERMAN K.
XX (MALY/) MALYANKAR U M.
XX (ORTT/) ORT T.
XX (GORM/) GORMAN L.
XX (ZERH/) ZERHUSEN B D.
XX (ANDE/) ANDERSON D W.
XX (ZHON/) ZHONG M.

```

```

PA (CATT/) CATTERTON E.
PA (JIWW/) JI W.
PA (MILL/) MILLER C E.
PA (RAST/) RASTELLI L.
PA (STON/) STONE D J.
PA (PENNA/) PENNA C E A.
PA (SHEN/) SHENOY S G.
PA (SHIM/) SHIMKETS R A.
PA (ROTH/) ROTHENBERG M E.
PA (LEAC/) LEACH M D.
PA (AGEE/) AGEE M L.
PA (BERG/) BERGHS C.
PA (DIP/) DIPIPPO V A.
PA (EISE/) EISEN A.
PA (GANG/) GANGOLLI E A.
PA (RIEG/) RIEGER D K.
PA (SPAD/) SPADERNA S K.
XX
XX Smuthson G, Millet I, Peyman JA, Kekuda R, Ju J, Li L, Guo X;
XX Patuturajan M, Spytek KA, Edinger SR, Ellerman K, Malyankar UM;
XX Ort T, Gorman L, Zerhusen BD, Anderson DW, Zhong M, Catterton E;
XX Ji W, Miller CE, Rastelli L, Stone DJ, Pena CE, Shenoy SG;
XX Shimkets RA, Rothenberg ME, Leach MD, Agee ML, Bergths C, Dipippo VA;
XX Eisen A, Gangolli EA, Rieger DK, Spaderna SK;
XX
XX WPI; 2004-213931/20.
XX N-PSDB; ADN62806.
XX
XX Isolated NOVA polypeptides and nucleic acids, useful for preventing,
XX diagnosing and treating e.g. cancer, diabetes and Alzheimer's disease.
XX
XX Example 54; Page; 395DP; English.
XX
XX The invention relates to isolated NOVA polypeptides and polynucleotides.
XX NOVA polypeptides and polynucleotides are used to prevent, diagnose or
XX treat a medical condition in human related to the aberrant expression and
XX activity of NOVA polypeptides. For example, NOVA polypeptides and
XX polynucleotides may be used to treat disorders associated with decreased
XX expression or activity of NOVA by supplementing the patient our
XX production or to rectify mutations. Conversely, antisense NA molecules
XX may be administered to down regulate expression of NOVA polypeptides by
XX binding with the cells own genes and preventing their expression. NOVA
XX polynucleotides and complementary sequences may also be used as DNA
XX probes in diagnostic assays to detect and quantify the presence of
XX similar sequences in samples, and so which patients may be in need of
XX restorative therapy. NOVA polypeptides may also be used as antigens in
XX the production of antibodies and in assays to identify modulators
XX (agonists and antagonists) of the expression and activity of NOVA. The
XX anti-NOVA polypeptide antibodies, agonists and antagonists may also be
XX used to modulate NOVA polynucleotide expression and activity of NOVA
XX polypeptides. The anti-NOVA polypeptide antibodies may also be used as
XX diagnostic agents for detecting the presence of NOVA in samples. NOVA
XX polypeptides and polynucleotides may be used in this way to prevent,
XX diagnose and treat: metabolic disorders, diabetes, obesity, infectious
XX disease, anorexia, cancer, cancer-associated cachexia, neurodegenerative
XX disorders, Alzheimer's Disease, Parkinson's Disorder, immune disorders,
XX haematopoietic disorders, and the various dyslipidaemias, metabolic
XX disturbances associated with obesity, the metabolic syndrome X and
XX wasting disorders associated with chronic diseases and various cancers.
XX They may also be used as antibacterial agents. The present sequence
XX represents the amino acid sequence of a human NOVA protein. Note the
XX present sequence is not shown in the specification but was created by the
XX indexer using the information given in example 54.
XX
XX Sequence 2053 AA:
XX
XX Query Match 99.2%; Score 10412.5; DB 8; Length 2053;
XX Best Local Similarity 99.4%; Pred. No. 0;
XX Matches 2042; Conservative 2; Mismatches 9; Indels 1; Gaps 1;
XX
XX 1 MLKFKYCARNPDLNAGAAEPFLASASRLNLFQCKPPPTQOQMSPLSRGIIALPLFE 60
XX 1 MLKFKYCARNPDLNAGAAEPFLASASRLNLFQCKPPPTQOQMSPLSRGIIALPLFE 60

```


QY 61 ECSOPALMKIYGVNFWFKYISDTTAELOELOPSADCFVRSIVGCGFAEVQVVERKATG 120
 DB 61 ECSOPALMKIYGVNFWFKYISDTTAELOELOPSADCFVRSIVGCGFAEVQVVERKATG 120
 QY 121 DIYAMKWKKKALLAOBOVSFFEEERNLSRSTSWIQLQOAFODKXHLVWNEYQGG 180
 DB 121 DIYAMKWKKKALLAOBOVSFFEEERNLSRSTSWIQLQOAFODKXHLVWNEYQGG 180
 QY 181 DLILSLNRYEDQDLENILQFYLAELILAVHSVHLMGVYHRDIKPENILVDRYGHKLVD 240
 DB 181 DLILSLNRYEDQDLENILQFYLAELILAVHSVHLMGVYHRDIKPENILVDRYGHKLVD 240
 QY 241 GSAAKMSNNKVNAKLPITGPDYAPVLTVMNGDGKGTGYGLDCMWSVGIAYEMTYGR 300
 DB 241 GSAAKMSNNKVNAKLPITGPDYAPVLTVMNGDGKGTGYGLDCMWSVGIAYEMTYGR 300
 QY 301 SPFAGTSARTFNINMFQRLKFPDDPKVSDFLDLQSLCCQKELKFBGLCCHFPF 360
 DB 301 SPFAGTSARTFNINMFQRLKFPDDPKVSDFLDLQSLCCQKELKFBGLCCHFPF 360
 QY 361 SKIDMNNIRNPPFVPTLKSDDTSNFBEPKNSWSSPCQJSPSGFSGEELPFVGS 420
 DB 361 SKIDMNNIRNPPFVPTLKSDDTSNFBEPKNSWSSPCQJSPSGFSGEELPFVGS 420
 QY 421 YSKALGILGRSESVSGIDSPAKTSSMEKULLIKSKELQDSQDKCHKXQDMTRYLHRRVS 480
 DB 421 YSKALGILGRSESVSGIDSPAKTSSMEKULLIKSKELQDSQDKCHKXQDMTRYLHRRVS 480
 QY 481 EVELVLSQKVEYELKASETORSLIFODLATYTCSSSLKSLSEQARMEVSQGDYALQILH 540
 DB 481 EVELVLSQKVEYELKASETORSLIFODLATYTCSSSLKSLSEQARMEVSQGDYALQILH 540
 QY 541 DIREQSRKLOEIKOEYOAOVEEMRLMNOLEBDLVASRRSDIYSESELRESRLAAEFK 600
 DB 541 DIREQSRKLOEIKOEYOAOVEEMRLMNOLEBDLVASRRSDIYSESELRESRLAAEFK 600
 QY 601 RKATECOHKLKAKDOQKPEVGEYAKLEKINAEOQLKQELQEKLEKAVKASTATELLO 660
 DB 601 RKATECOHKLKAKDOQKPEVGEYAKLEKINAEOQLKQELQEKLEKAVKASTATELLO 660
 QY 661 NIROAKRARELEKLONRDSSGICRKKVLEABERHNSLENKYKLETERRENNRLKDD 720
 DB 661 NIROAKRARELEKLONRDSSGICRKKVLEABERHNSLENKYKLETERRENNRLKDD 720
 QY 721 IOTKSQOIQQMADKILEEKHREAOVSAGHLVEHLKOKBOHYBEKLVLDNOIKKDLAD 780
 DB 721 IOTKSQOIQQMADKILEEKHREAOVSAGHLVEHLKOKBOHYBEKLVLDNOIKKDLAD 780
 QY 781 KETLENNMOHBEBAHEKGLISBOKAMINAMDSKIRSLBORIYELSEANKLANSSILFT 840
 DB 781 KETLENNMOHBEBAHEKGLISBOKAMINAMDSKIRSLBORIYELSEANKLANSSILFT 840
 QY 841 QRMNKAQEBEIMSELRQOKFYLETQAGKLEAQRNLEQOLEKISHQDSDSKRRLLEETRL 900
 DB 841 QRMNKAQEBEIMSELRQOKFYLETQAGKLEAQRNLEQOLEKISHQDSDSKRRLLEETRL 900
 QY 901 REVSLEHEBOKLEKROLTELQSLORRESQLTALQARABALBEOQLQAKTELEETTAEA 960
 DB 901 REVSLEHEBOKLEKROLTELQSLORRESQLTALQARABALBEOQLQAKTELEETTAEA 960
 QY 961 EEBEIOALTARHDEIQRFKDALRNSCTVITDLEBOLNLTEDNAELNNQNFYLSKQLEBAS 1020
 DB 961 EEBEIOALTARHDEIQRFKDALRNSCTVITDLEBOLNLTEDNAELNNQNFYLSKQLEBAS 1020
 QY 1021 GANDEIVOLRSEVDHLREITEREMOULTSOKQVNEALKTCTMLBEOVMDLEBALNDELLE 1080
 DB 1021 GANDEIVOLRSEVDHLREITEREMOULTSOKQVNEALKTCTMLBEOVMDLEBALNDELLE 1080
 QY 1081 KEROMEAMRSVTLGDEKSOFEGRVLELQMLDTEKOSBARADORTTESROVVELAVKXHA 1140
 DB 1081 KEROMEAMRSVTLGDEKSOFEGRVLELQMLDTEKOSBARADORTTESROVVELAVKXHA 1140
 QY 1080 KEROMEAMRSVTLGDEKSOFEGRVLELQMLDTEKOSBARADORTTESROVVELAVKXHA 1139
 DB 1080 KEROMEAMRSVTLGDEKSOFEGRVLELQMLDTEKOSBARADORTTESROVVELAVKXHA 1139

QY 1141 EILALQALKEOKUKAASLSKINDLEKXHAMLENNARSLOOKLETERBELKORLLEBOAK 1200
 DB 1141 EILALQALKEOKUKAASLSKINDLEKXHAMLENNARSLOOKLETERBELKORLLEBOAK 1199
 QY 1201 LQOQMDLOKXNIFPLTQLOFALRADILKTERBSDLFQULENIOVLVYSEHKYKMGSTISQ 1260
 DB 1201 LQOQMDLOKXNIFPLTQLOFALRADILKTERBSDLFQULENIOVLVYSEHKYKMGSTISQ 1259
 QY 1261 QTKLIDFLQAMMDPAKKKKVPLQYNELKALBEXKARCALBERLOQTRIELRSAREBA 1320
 DB 1261 QTKLIDFLQAMMDPAKKKKVPLQYNELKALBEXKARCALBERLOQTRIELRSAREBA 1319
 QY 1321 AHRKATDHPHSTPATARQOIJAMSAIVSPHQSAMSILAPSSRRKESSTPEEFSSRL 1380
 DB 1321 AHRKATDHPHSTPATARQOIJAMSAIVSPHQSAMSILAPSSRRKESSTPEEFSSRL 1379
 QY 1381 KERHNNHNPHRFNGLNMRATKCAVCLDTYHFGQASCTLECCOVWCHKXCTCLPATGGL 1440
 DB 1381 KERHNNHNPHRFNGLNMRATKCAVCLDTYHFGQASCTLECCOVWCHKXCTCLPATGGL 1439
 QY 1441 PAEYATHTPEAFCDKKNNSPGLQTKPESSSLHLEGMKVPANNKRGQOQMDPKTYVLEGS 1500
 DB 1441 PAEYATHTPEAFCDKKNNSPGLQTKPESSSLHLEGMKVPANNKRGQOQMDPKTYVLEGS 1499
 QY 1501 KVLIDNEAREBARAGRPVEBEFLCLPDGDVSJHGAVGASELANTAKADVPYILKXESHPT 1560
 DB 1501 KVLIDNEAREBARAGRPVEBEFLCLPDGDVSJHGAVGASELANTAKADVPYILKXESHPT 1559
 QY 1561 TCWPGRTLYLLABSPFPKORVNTLBSVAVAGRSREKAEADAKLNGSLKLEGGDRILD 1620
 DB 1561 TCWPGRTLYLLABSPFPKORVNTLBSVAVAGRSREKAEADAKLNGSLKLEGGDRILD 1619
 QY 1621 NCTLPSPSDQVVLVGTSEGLYALNVLKNSLTHVIGIGAFQIYIILKOLEKLMTAGEBERA 1680
 DB 1621 NCTLPSPSDQVVLVGTSEGLYALNVLKNSLTHVIGIGAFQIYIILKOLEKLMTAGEBERA 1679
 QY 1681 LCIVDVKKVQSLAQSHLPAPDIPSIPIPEAVKGCFLGAKIENGGLCTCAMPSKVYL 1740
 DB 1681 LCIVDVKKVQSLAQSHLPAPDIPSIPIPEAVKGCFLGAKIENGGLCTCAMPSKVYL 1739
 QY 1741 RYENLSKYCIKKEIETSEPOSCITHFTNYSLILGTNKFYEDMKQYTLBEFLDKNDHSLA 1800
 DB 1741 RYENLSKYCIKKEIETSEPOSCITHFTNYSLILGTNKFYEDMKQYTLBEFLDKNDHSLA 1799
 QY 1801 PAVFAASSNPVSIVOVNSAGOREEYLLCFHEGFVVDSTGRRSKTDDLKMSRLPLAFA 1860
 DB 1801 PAVFAASSNPVSIVOVNSAGOREEYLLCFHEGFVVDSTGRRSKTDDLKMSRLPLAFA 1859
 QY 1861 YREBYLFTVTHFNSLEVEIQARSSAGTPARAYLDIPMPRYLGPAISSGAYLAASYODKL 1920
 DB 1861 YREBYLFTVTHFNSLEVEIQARSSAGTPARAYLDIPMPRYLGPAISSGAYLAASYODKL 1919
 QY 1921 RVICCKGNLVKESGTBEHHRGPOSTSRSPNKGPTVNEHTIKRVASSPAPPEGSHPREP 1980
 DB 1921 RVICCKGNLVKESGTBEHHRGPOSTSRSPNKGPTVNEHTIKRVASSPAPPEGSHPREP 1979
 QY 1981 STPHRYREGRTBLRDSKPGRLERKSPGRILSTREBRSPARVFEDESSRGLPLAGAVRT 2040
 DB 1981 STPHRYREGRTBLRDSKPGRLERKSPGRILSTREBRSPARVFEDESSRGLPLAGAVRT 2039
 QY 2041 PLSQVNCVMDOSSV 2054
 DB 2041 PLSQVNCVMDOSSV 2053

RESULT 12
 ADN62819
 ID ADN62819 standard; protein; 2066 AA.
 XX
 AC ADN62819;
 XX
 DT 01-JUL-2004 (first entry)
 XX

DE Human NOV1g.
 XX
 KW human; NOVX; metabolic disorder; diabetes; obesity; infectious disease;
 KW anorexia; cancer; cancer-associated cachexia; neurodegenerative disorder;
 KW Alzheimer's disease; Parkinson's disease; immune disorder;
 KW hematoepietic disorder; dyslipidaemia; metabolic syndrome X;
 KW wasting disorder.
 OS Homo sapiens.
 XX
 PN US200403823-A1.
 XX
 PD 26-FEB-2004.
 XX
 PF 01-OCT-2002; 2002US-00262511.
 XX
 PR 02-OCT-2001; 2001US-0326483P.
 PR 05-OCT-2001; 2001US-0327443P.
 PR 09-OCT-2001; 2001US-0327449P.
 PR 09-OCT-2001; 2001US-0327917P.
 PR 09-OCT-2001; 2001US-0328029P.
 PR 09-OCT-2001; 2001US-0328044P.
 PR 09-OCT-2001; 2001US-0328056P.
 PR 12-OCT-2001; 2001US-0328849P.
 PR 15-OCT-2001; 2001US-0329414P.
 PR 17-OCT-2001; 2001US-0330142P.
 PR 18-OCT-2001; 2001US-0330309P.
 PR 22-OCT-2001; 2001US-0341058P.
 PR 24-OCT-2001; 2001US-0339266P.
 PR 24-OCT-2001; 2001US-0343629P.
 PR 29-OCT-2001; 2001US-0349575P.
 PR 01-NOV-2001; 2001US-0346357P.
 PR 17-APR-2002; 2002US-0373260P.
 PR 19-APR-2002; 2002US-0373815P.
 PR 19-APR-2002; 2002US-0373826P.
 PR 19-APR-2002; 2002US-0373884P.
 PR 19-APR-2002; 2002US-0374977P.
 PR 22-APR-2002; 2002US-0381037P.
 PR 16-MAY-2002; 2002US-0381038P.
 PR 16-MAY-2002; 2002US-0381042P.
 PR 17-MAY-2002; 2002US-0381642P.
 PR 28-MAY-2002; 2002US-0383656P.
 PR 29-MAY-2002; 2002US-0383831P.
 PR 25-JUN-2002; 2002US-0391335P.
 XX
 PA (SMIT/) SMITHSON G.
 PA (MILL/) MILLET I.
 PA (PEYM/) PEYMAN J A.
 PA (KEKU/) KEKUDA R.
 PA (JUUJ/) JU J.
 PA (LILL/) LI L.
 PA (GUOX/) GUO X.
 PA (PATT/) PATTURAJAN M.
 PA (SPYT/) SPYTEK K A.
 PA (EDIN/) EDINGER S R.
 PA (ELLE/) ELLERMAN K.
 PA (MALY/) MALYANKAR V M.
 PA (ORTT/) ORT T.
 PA (GORM/) GORMAN L.
 PA (ZERR/) ZERRHUSEN B D.
 PA (ANDR/) ANDERSON D W.
 PA (ZHON/) ZHONG M.
 PA (CATT/) CATTERTON E.
 PA (JIMM/) JI W.
 PA (MILL/) MILLER C E.
 PA (RAST/) RASTTELLI L.
 PA (STON/) STONE D J.
 PA (PENA/) PENNA C B A.
 PA (SHEN/) SHENOY S G.
 PA (SHIM/) SHIMKETS R A.
 PA (ROTH/) ROTHENBERG M E.
 PA (LEAC/) LEACH M D.

PA (AGEE/) AGEER M L.
 PA (BERG/) BERGHS C.
 PA (DIP/) DIPOLPO V A.
 PA (EISE/) EISEN A.
 PA (GANG/) GANGOLLI E A.
 PA (RIEG/) RIEGER D K.
 PA (SPAD/) SPADERNA S K.
 XX
 PI Smithson G, Millet I, Peyman JA, Kekuda R, Ju J, Li L, Guo X;
 PI Patutturajan M, Spytek KA, Edinger SR, Ellerman K, Malyankar VM,
 PI Ort T, Gorman L, Zerrhuseen BD, Anderson DW, Zhong M, Catterton E;
 PI Ji W, Miller CE, Rastelli L, Stone DJ, Penna CE, Sheno S, Shimk R,
 PI Shimkets RA, Rothenberg ME, Leach MD, Agee ML, Berghs C, Dipolpo VA,
 PI Eisen A, Gangolli EA, Rieger DK, Spaderna SK;
 XX
 DR WPI; 2004-213931/20.
 DR N-PSDB; ADN62818.
 XX
 PT Isolated NOVX polypeptides and nucleic acids, useful for preventing,
 PT diagnosing and treating e.g. cancer, diabetes and Alzheimer's disease.
 XX
 PS Claim 1; SEQ ID NO 14; 395pp; English.
 XX
 CC The invention relates to isolated NOVX polypeptides and polynucleotides.
 CC NOVX polypeptides and polynucleotides are used to prevent, diagnose or
 CC treat a medical condition in human related to the aberrant expression and
 CC activity of NOVX polypeptides. For example, NOVX polypeptides and
 CC polynucleotides may be used to treat disorders associated with decreased
 CC expression or activity of NOVX by supplementing the patient our
 CC production or to rectify mutations. Conversely, antisense NA molecules
 CC may be administered to down regulate expression of NOVX polypeptides by
 CC binding with the cells own genes and preventing their expression. NOVX
 CC polynucleotides and complementary sequences may also be used as DNA
 CC probes in diagnostic assays to detect and quantitate the presence of
 CC similar sequences in samples, and so which patients may be in need of
 CC restorative therapy. NOVX polypeptides may also be used as antigens in
 CC the production of antibodies and in assays to identify modulators
 CC (agonists and antagonists) of the expression and activity of NOVX. The
 CC anti-NOVX polypeptide antibodies, agonists and antagonists may also be
 CC used to modulate NOVX polynucleotide expression and activity of NOVX
 CC polypeptides. The anti-NOVX polypeptide antibodies may also be used as
 CC diagnostic agents for detecting the presence of NOVX in samples. NOVX
 CC polypeptides and polynucleotides may be used in this way to prevent,
 CC diagnose and treat: metabolic disorders, diabetes, obesity, infectious
 CC disease, anorexia, cancer, cancer-associated cachexia, neurodegenerative
 CC disorders, Alzheimer's Disease, Parkinson's Disorder, immune disorders,
 CC haematopoietic disorders, and the various dyslipidaemias, metabolic
 CC disturbances associated with obesity, the metabolic syndrome X and
 CC wasting disorders associated with chronic diseases and various cancers.
 CC They may also be used as antibacterial agents. The present sequence
 CC represents the amino acid sequence of a human NOVX protein.
 XX
 SQ Sequence 2066 AA;
 Query Match 99.2%; Score 10407.5; DB 8; Length 2066;
 Best Local Similarity 99.5%; Pred. No. 0;
 Matches 2041; Conservative 2; Mismatches 8; Indels 1; Gaps 1;
 QY 1 MLKFKYGARNPIDAGAAEPPIASRASRLNTFFQKPPFMTQOQMSPLISREGIDALFVLF 60
 DB 1 MLKFKYGARNPIDAGAAEPPIASRASRLNTFFQKPPFMTQOQMSPLISREGIDALFVLF 60
 QY 61 ECGQPALMKIKHVSNTVRKYSPTIAELOIOPSAKPFVRSLVCGGHFAVQVVRKATG 120
 DB 61 ECGQPALMKIKHVSNTVRKYSPTIAELOIOPSAKPFVRSLVCGGHFAVQVVRKATG 120
 QY 121 DIYAMKVMKKKALLAEOVSFFEEERNILSRSTSPWIPOLQYAFODKNHLYLWMEYOPGG 180
 DB 121 DIYAMKVMKKKALLAEOVSFFEEERNILSRSTSPWIPOLQYAFODKNHLYLWMEYOPGG 180
 QY 181 DLLSLNRYEDQDENLIOFYLAELILAVSHVLMGVHRHDIPENILVDRTGHIKLVDF 240
 DB 181 DLLSLNRYEDQDENLIOFYLAELILAVSHVLMGVHRHDIPENILVDRTGHIKLVDF 240

QY 241 GSAKXNSNMVNAKLPICGPDYMAPEVLTVMNGDGKTYGLDCDDWMSVGVIAVEATYGR 300
 DB 241 GSAKXNSNMK-VNAKLPICGPDYMAPEVLTVMNGDGKTYGLDCDDWMSVGVIAVEATYGR 299
 QY 301 SPFAEGTSARTFNNIMNFORFLKFPDDPKVSSDLDLIQSLCGQKRLKEGKCPHPF 360
 DB 300 SPFAEGTSARTFNNIMNFORFLKFPDDPKVSSDLDLIQSLCGQKRLKEGKCPHPF 359
 QY 361 SKIDMNNIRNSPPFPVPTLSDDDTSNFDEBEKNSWVSSPFCQSPSGFSGEELPFVGS 420
 DB 360 SKIDMNNIRNSPPFPVPTLSDDDTSNFDEBEKNSWVSSPFCQSPSGFSGEELPFVGS 419
 QY 421 YSKRLGLIGSSESVVSGLDSPAKTSMEKLLIYSKELQDSODYKCHMEQMTLHRVS 480
 DB 420 YSKRLGLIGSSESVVSGLDSPAKTSMEKLLIYSKELQDSODYKCHMEQMTLHRVS 479
 QY 481 EYEAVALSOKVEYELKASTORSILFODLATYTCSSSLKRSLEQARMESQEDDQALDHL 540
 DB 480 EYEAVALSOKVEYELKASTORSILFODLATYTCSSSLKRSLEQARMESQEDDQALDHL 539
 QY 541 DIREOSKLOEIKOEYOAOVEEMRLMNNQLEEDLVARRRSDIYSESELSESLAAEFK 600
 DB 540 DIREOSKLOEIKOEYOAOVEEMRLMNNQLEEDLVARRRSDIYSESELSESLAAEFK 599
 QY 601 RKATECOHKLKAKDQKPEVGEYAKLEKINAEOQLKIQELQELKAVASTEATSLQ 660
 DB 600 RKATECOHKLKAKDQKPEVGEYAKLEKINAEOQLKIQELQELKAVASTEATSLQ 659
 QY 661 NIIOAKRARELEKONREDSSEGIIRKULVEAERHSLEENKYLFTMERRENRLKDD 720
 DB 660 NIIOAKRARELEKONREDSSEGIIRKULVEAERHSLEENKYLFTMERRENRLKDD 719
 QY 721 IQTSQOIQOMADKILLEEKHRAOYSAQHLVHLKQKQHYEKLKVLNOKKDLAD 780
 DB 720 IQTSQOIQOMADKILLEEKHRAOYSAQHLVHLKQKQHYEKLKVLNOKKDLAD 779
 QY 781 KETLENNMOHSEEAHEKGISLEQKAMINAMDSKIRSLFORIVELSEANKLAANSLSFT 840
 DB 780 KETLENNMOHSEEAHEKGISLEQKAMINAMDSKIRSLFORIVELSEANKLAANSLSFT 839
 QY 841 QRMKAOEMISLIRQOKFLFETOAGKLEAONRLLEOKELSHQODSDKNRLLEETRL 900
 DB 840 QRMKAOEMISLIRQOKFLFETOAGKLEAONRLLEOKELSHQODSDKNRLLEETRL 899
 QY 901 REVSLEHEOKLEIKROLTELQSLQERESQTLQARALAESQLOAKTELETTAA 960
 DB 900 REVSLEHEOKLEIKROLTELQSLQERESQTLQARALAESQLOAKTELETTAA 959
 QY 961 EBEIOALTARDEIORKFDALRNSCTVITDLEBQNLQTEDNAELNNQNFYLSQDLDEAS 1020
 DB 960 EBEIOALTARDEIORKFDALRNSCTVITDLEBQNLQTEDNAELNNQNFYLSQDLDEAS 1019
 QY 1021 GANDEIYOLRSEVDHLREITTEBMOULTSQOKYMEALKTTCTMLEBOVMLEALNDELE 1080
 DB 1020 GANDEIYOLRSEVDHLREITTEBMOULTSQOKYMEALKTTCTMLEBOVMLEALNDELE 1079
 QY 1081 KEROMEAMRSVILGDEKSOFCRAVELQRMIDTEKOSARADORTTESROVVELAVKEHKA 1140
 DB 1080 KEROMEAMRSVILGDEKSOFCRAVELQRMIDTEKOSARADORTTESROVVELAVKEHKA 1139
 QY 1141 EIALAQALKEOKLKESLSDKLNDELEKHAMLENNARSIOQLKLETRRELKORLLEQAK 1200
 DB 1140 EIALAQALKEOKLKESLSDKLNDELEKHAMLENNARSIOQLKLETRRELKORLLEQAK 1199
 QY 1201 LOOQMDLOKNIIFRLTQLOALDRADLKTESDLEYOLENTQVLSHEKVMKEGTISQ 1260
 DB 1200 LOOQMDLOKNIIFRLTQLOALDRADLKTESDLEYOLENTQVLSHEKVMKEGTISQ 1259
 QY 1261 QTKLIDFLQAKMDOPAKKKVPLQVNETKLALKEKAKCALEBALQKTRIELRSAREEA 1320
 DB 1260 QTKLIDFLQAKMDOPAKKKVPLQVNETKLALKEKAKCALEBALQKTRIELRSAREEA 1319

QY 1321 AHRKATDHPSPPTARAOQIAMSATVRSPEHQSAMSILAPSSRKESTPEEFSTR 1380
 DB 1320 AHRKATDHPSPPTARAOQIAMSATVRSPEHQSAMSILAPSSRKESTPEEFSTR 1379
 QY 1381 KERMEHNI PHRFVNGLMARATKCAVCDTVHFGROAKCLBQVCHPCKSTCIPATG 1440
 DB 1380 KERMEHNI PHRFVNGLMARATKCAVCDTVHFGROAKCLBQVCHPCKSTCIPATG 1439
 QY 1441 PAEYATHTFAFCRDKNSPGLQTEPSSSLHLEGMMKVRNNRGOQMDRKYIVLEGS 1500
 DB 1440 PAEYATHTFAFCRDKNSPGLQTEPSSSLHLEGMMKVRNNRGOQMDRKYIVLEGS 1499
 QY 1501 KVLIVNDEARBAQORPVEEFELCLPDGDVSIHGAVGSSELANTKADAVPYILKMEHPHT 1559
 DB 1500 KVLIVNDEARBAQORPVEEFELCLPDGDVSIHGAVGSSELANTKADAVPYILKMEHPHT 1559
 QY 1561 TCMFGRTLYLLAPSPKORWVTALESVAVGGRVSRKAEADAKLNSLILKLEGDRLD 1620
 DB 1560 TCMFGRTLYLLAPSPKORWVTALESVAVGGRVSRKAEADAKLNSLILKLEGDRLD 1619
 QY 1621 MNCGLPSPDQVVLVGTBEGYALNVLKNSLTHVPGI GAVFOIYILKOLEKLMITAGEERA 1680
 DB 1620 MNCGLPSPDQVVLVGTBEGYALNVLKNSLTHVPGI GAVFOIYILKOLEKLMITAGEERA 1679
 QY 1681 LCLVDVKKVQSLAQSHLPAQPDISPNIFPAVVGCHLFGAGKLENGLICAMAPSKVIL 1740
 DB 1680 LCLVDVKKVQSLAQSHLPAQPDISPNIFPAVVGCHLFGAGKLENGLICAMAPSKVIL 1739
 QY 1741 RYNNENLSKYCIKREIETSEBSCCHFTNYSILGNTNFYEIDMKQYTLSEFLDKNDHSLA 1800
 DB 1740 RYNNENLSKYCIKREIETSEBSCCHFTNYSILGNTNFYEIDMKQYTLSEFLDKNDHSLA 1799
 QY 1801 PAVFAASNSFPVSIYOVNSAGOREEYLCFHEFGVFDVSGRSRTDDLKMSRLPLAFA 1860
 DB 1800 PAVFAASNSFPVSIYOVNSAGOREEYLCFHEFGVFDVSGRSRTDDLKMSRLPLAFA 1859
 QY 1861 YREBYLFTVTHFNSLEVEIIOARSSAGTPARAYLDIPNRYLGPALISSGAYLASSYODKL 1920
 DB 1860 YREBYLFTVTHFNSLEVEIIOARSSAGTPARAYLDIPNRYLGPALISSGAYLASSYODKL 1919
 QY 1921 RVLCCKGNLVKESGTEHHRGPTSRSSPNKRGPPTVNEHTTKRVASSPAPEGSHPREP 1980
 DB 1920 RVLCCKGNLVKESGTEHHRGPTSRSSPNKRGPPTVNEHTTKRVASSPAPEGSHPREP 1979
 QY 1981 STPHRYEGRTELRDOKSPGRPLEREKSPGRILSTRERESPALFEDSSGRPLPAGAVRT 2040
 DB 1980 STPHRYEGRTELRDOKSPGRPLEREKSPGRILSTRERESPALFEDSSGRPLPAGAVRT 2039
 QY 2041 PLSQVNNKWDQS 2052
 DB 2040 PLSQVNNKVRQHS 2051

RESULT 13
 ABP97683
 ID ABP97683 standard; protein; 2055 AA.
 AC ABP97683;
 DT 16-MAY-2003 (first entry)
 XX Polypeptide similar to citron rho/rac-interacting kinase-short kinase.
 XX Human; citron rho/rac-interacting kinase-short kinase; obesity;
 KW chronic obstructive pulmonary disease; hypertension; diabetes;
 KW coronary artery disease; hyperlipidemia; stroke; gallbladder disease;
 KW gout; osteoarthritis; sleep apnea; cancer; thrombotic disease;
 KW polycystic ovarian syndrome; fertility; depression.
 XX Homo sapiens.
 OS Homo sapiens.
 XX MO2003004629-A2.
 XX

Query Match	Best Local Similarity	96.3%	Score 10109.5	DB 6	Length 2055
Matches 1974	Conservative 40	Mismatches 39	Indels 3	Gaps 2	
QY	1	MLKFKYGRNPIDAGAAEPDASRASRLNLFQCKPEPMTQOMSPLSREGIIDLFLVLE	60		
DB	1	MLKFKYGRNPPEAASEPIASRASRLNLFQCKPEPMTQOMSPLSREGIIDLFLVLE	60		
QY	61	ECGQALMKIKHVSFVRKRYSTTIAELQPSADFEVRSIVGCGHFAEYQVVERKATG	120		
DB	61	ECGQALMKIKHVSFVRKRYSTTIAELQPSADFEVRSIVGCGHFAEYQVVERKATG	120		
QY	121	DIYAKVWKKKALLAQEOVSFFEEERNILSRSTFWIPOLQYAFODKNLVLVMEYOPGG	180		
DB	121	DYVANKIMKKKALLAQEOVSFFEEERNILSRSTFWIPQLQYAFODKNLVLVMEYOPGG	180		
QY	181	DLISLIRNEYDQIDENLIQFYLAELILAVHSVHLMGYHARDIKPENILVDRGHIKLVD	240		
DB	181	DFLSLIRNEYDQIDENLIQFYLAELILAVHSVHLMGYHARDIKPENILVDRGHIKLVD	240		
QY	241	GSAAGKNSNKNVNAALPIGTPTPYMAPEVLTWNNGGKGYGLDCPMWSGVAIAYETVGR	300		
DB	241	GSAAGKNSNKNVNAALPIGTPTPYMAPEVLTWNNEGRGTYGLDCPMWSGVAIAYETVGR	299		
QY	301	SPFAEGTSARTPTNNIMNFORFLKFPDDPKVSSDFIDLQSLLCGQKERYLKEFGELCHPFF	360		
DB	300	TPFTEGTSARTPTNNIMNFORFLKFPDDPKVSSDFIDLQSLLCGQKERYLKEFGELCHPFF	359		
QY	361	SKIDNNIRNSPPPVPTLKSDDDTSNPFPEPKNSWSSPCQLSPSGSGEELPVGFS	420		
DB	360	ARTDNNIRNSPPPVPTLKSDDDTSNPFPEPKNSWAFILCVPAEPLAFSGEELPVGFS	419		
QY	421	YSKALGILGRSSSVSGDPSPAKTSMSKKLLIKSKELDSDCKHGMQEWTRILHRRYS	480		
DB	420	YSKALGILGRSSSVSGDPSPAKTSMSKKLLIKSKELDSDCKHGMQEWTRILHRRYS	479		

QY	481	IWEAVLSCKEVELKASEFORSLEBODATYITTECSSIKRSLSEQARMEVSQEDDVALQJLH	540
DB	480	EEVAVLSQKEVELKASEFORSLEBODATYITTECSSIKRSLSEQARMEVSQEDDVALQJLH	539
QY	541	DIREOSRLOJIKEOEYOAQVEENRLMNNQLEEDVSNARRSDYSESELRSRLAAEFK	600
DB	540	DIREOSRLOJIKEOEYOAQVEENRLMNNQLEEDVSNARRSDYSESELRSRLAAEFK	599
QY	601	RKATECQHKLLKADQCKPEVGEYAKI, EKINAEQOLKI QELQELKRAVKASTATELLQ	660
DB	600	RKANECHKLMKADQCKPEVGEYAKI, EKINAEQOLKI QELQELKRAVKASTATELLQ	659
QY	661	NIRAKRABERLEKLNREDSGGIKKIVAEERBSHLENKYKRL, ETMRERREMLKD	720
DB	660	NIRAKRABERLEKLNREDSGGIKKIVAEERBSHLENKYKRL, ETMRERREMLKD	719
QY	721	IQTSQOIQQVADKIL, ELEEKHRAOVSQAH, EYHLNKOQEOHYEKKI KVLDNOIKKDLAD	780
DB	720	IQTSQOIQQVADKIL, ELEEKHRAOVSQAH, EYHLNKOQEOHYEKKI KVLDNOIKKDLAD	779
QY	781	KETLENNMQRHEBAHEKGLTSLSQKAMINANDSKTNSLEORIVEL, SEANKLANSS, LFT	840
DB	780	KESLENNMQRHEBAHEKGLTSLSQKAMINANDSKTNSLEORIVEL, SEANKLANSS, LFT	839
QY	841	QRNKAQEMMISELRQCKFYLETQAGLEQONRKL, EQL, EKISHODSDKXRL, ELEFRL	900
DB	840	QRNKAQEMMISELRQCKFYLETQAGLEQONRKL, EQL, EKISHODSDKXRL, ELEFRL	899
QY	901	REVSLEHEBQELKRLQUTELQTSLOERESQTLQAPARAL, EBSQLRQAKTELEETTAEA	960
DB	900	REVSLEHEBQELKRLQUTELQTSLOERESQTLQAPARAL, EBSQLRQAKTELEETTAEA	959
QY	961	EBEIQALTAHDEIQRRKFDALRNSCTVI, TDLEEQLNQ, TEDNAELNNOFYL, SQDLBAS	1020
DB	960	EBEIQALTAHDEIQRRKFDALRNSCTVI, TDLEEQLNQ, TEDNAELNNOFYL, SQDLBAS	1019
QY	1021	GANDEIYQJNSEVDHRLREITEREMOLTSQQTMEAKTCTM, L, EBOVMDE, EALNDELLE	1080
DB	1020	GANDEIYQJNSEVDHRLREITEREMOLTSQQTMEAKTCTM, L, EBOVMDE, EALNDELLE	1079
QY	1081	KEROMEAMRSVLGDEKSOFECEVRELQRM, DTEKOSPAPADQRTTESQVVELAVKEHKA	1140
DB	1080	KEROMEAMRSVLGDEKSOFECEVRELQRM, DTEKOSPAPADQRTTESQVVELAVKEHKA	1139
QY	1141	EIHALQOALKEQKIKAESLSDKNDLDEKXAMLEMNARS, LQOKLETERELKORLL, EBOAK	1200
DB	1140	EIHALQOALKEQKIKAESLSDKNDLDEKXAMLEMNARS, LQOKLETERELKORLL, EBOAK	1199
QY	1201	LQOQMDLOKXNI, PULTOGLQELADRALDLKTERSDYLEQ, ENIOV, YSHEVKNMEGTT, SQ	1260
DB	1200	LQOQMDLOKXNI, PULTOGLQELADRALDLKTERSDYLEQ, ENIOV, YSHEVKNMEGTT, SQ	1259
QY	1261	QTKLIDFLQAMDP, PAKKKKVP, LYNEK, LAL, EKEKKARCALE, EALQKTR, IELTARSABEA	1320
DB	1260	QTKLIDFLQAMDP, PAKKKKVP, LYNEK, LAL, EKEKKARCALE, EALQKTR, IELTARSABEA	1319
QY	1321	AHRKATDHPHSTPATARQOJLAMS, AIVRSPEHOSAMS, L, APSSRRKESSTPEEFSRRL	1380
DB	1320	AHRKATDHPHSTPATARQOJLAMS, AIVRSPEHOSAMS, L, APSSRRKESSTPEEFSRRL	1379
QY	1381	KERMHNHNI, PHRFNVGLMNRATKCAV, CJD, TYHFGQAS, KCL, ECVQWCHPKCST, CL, PATCGL	1440
DB	1380	KERMHNHNI, PHRFNVGLMNRATKCAV, CJD, TYHFGQAS, KCL, ECVQWCHPKCST, CL, PATCGL	1439
QY	1441	PAEYATHTPTEA, FCGDKXNSPGLQIKESSS, L, HLEGNMKV, PPNNNRGGQGMKRTI, YLEGS	1500
DB	1440	PAEYATHTPTEA, FCGDKXNSPGLQIKESSS, L, HLEGNMKV, PPNNNRGGQGMKRTI, YLEGS	1499
QY	1501	KVL, LYNDNEARAGOR, PVEEFEL, CL, PDDQVSI, HGA, VNGASE, LANTAKADVPY, LITKMSHPHT	1560
DB	1500	KVL, LYNDNEARAGOR, PVEEFEL, CL, PDDQVSI, HGA, VNGASE, LANTAKADVPY, LITKMSHPHT	1559
QY	1561	TCWQGRTL, YLILPSP, PDKQRWVTAL, ESVAVAGRV, SRREKAEADAK, L, GNSLILKLEBGGDR, LD	1620

```

Db      1560 TCWGRITLYLAPEPDKQKRWVTLAESVAGARVREVAEDAKLGLSLKLEDDLD
Qy      1621 MNCITLPSDDQVVLVGTBEGVALNVLKNSLTHVPDVGAFQIYIIKDEKMLAGEERA
Db      1620 MNCITLPSDDQVVLVGTBEGVALNVLKNSLTHVPDVGAFQIYIIKDEKMLAGEERA
Qy      1681 LCLVDVKKVKQSLAQSHLPAPDIPSPNFEAVKGCHEFGAGKXENGLCTICAMPKSVIL
Db      1680 LCLVDVKKVKQSLAQSHLPAPDIPSPNFEAVKGCHEFGAGKXENGLCTICAMPKSVIL
Qy      1741 RYNNELSKYCKIKRKEIETSEPCSIHFTMYSLIGNKKEIYDMQYTLERFLDKNDHSLA
Db      1740 RYNNELSKYCKIKRKEIETSEPCSIHFTMYSLIGNKKEIYDMQYTLERFLDKNDHSLA
Qy      1801 PAVEAASNSFPVSIIVQVNSAGOREYLLCFHEFGVFDVSGRRSRTDLMKMSRLPLAFA
Db      1800 PAVEAASNSFPVSIIVQVNSAGOREYLLCFHEFGVFDVSGRRSRTDLMKMSRLPLAFA
Qy      1861 YREPYLFTVTHPNSLEVIEIQARSSAGTPARAYLDIPNRYGPAISGAYIYLAASYODKL
Db      1860 YREPYLFTVTHPNSLEVIEIQARSSAGTPARAYLDIPNRYGPAISGAYIYLAASYODKL
Qy      1921 RVICCKGNLVKESGTEHHRGSTSRSSPNKGPPTVYEHITKRYAASAPPEGSHPREP
Db      1920 RVICCKGNLVKESGTEHHRGSTSRSSPNKGPPTVYEHITKRYAASAPPEGSHPREP
Qy      1981 STPHRY--REGRTLRBDKSPGRPLEREKSPGRILSTFRERSPARLFEDESRRGLPLAGAV
Db      1980 STPHRY--REGRTLRBDKSPGRPLEREKSPGRILSTFRERSPARLFEDESRRGLPLAGAV
Qy      2039 RTPISQVNVKWDQSSV 2054
Db      2040 RTPISQVNVKWDQSSV 2055

RESULT 14
AAO26960 standard; protein; 2055 AA.
XX
AC      AAO26960;
XX
DE      01-MAY-2003 (first entry)
XX
KW      Human CRK related protein sequence, SEQ ID NO 3.
XX
KW      Anorectic; hypotensive; cardiatic; antilipemic; cerebroprotective;
KW      antitumor; osteopathic; antiarthritic; cytoskeletal; antidepressant;
KW      immunomodulator; antitumor; tranquilizer; antiparkinsonian; nootropic;
KW      neuroprotective; antiinflammatory; antidiabetic; analgesic;
KW      human citron rho/rac-interacting kinase; enzyme; CRK; ameliorating;
KW      obesity; comorbidity; cancer; anorexia; cachexia; bulimia;
KW      central nervous system disorder; chronic obstructive pulmonary disease;
KW      diabetes; pain.
XX
OS      Homo sapiens.
XX
PN      MO2003004523-A1.
XX
PD      16-JAN-2003.
XX
PF      28-JUN-2002; 2002MO-EP007156.
XX
PR      02-JUL-2001; 2001US-0301841P.
PR      11-DEC-2001; 2001US-0338651P.
PR      25-APR-2002; 2002US-0375014P.
XX
PA      (FARB ) BAYER AG.
XX
PI      Zhu Z;
XX
DR      WPI; 2003-221576/21.

```

```

PT      New human citron rho/rac-interacting kinase (CRK) polypeptide and
PT      polynucleotide, useful in preventing, ameliorating or treating diseases
PT      associated with human CRK dysfunction, e.g. obesity, diabetes or
PT      Alzheimer's disease.
XX
PS      Disclosure; Fig 3; 23pp; English.
XX
CC      The invention relates to an isolated polynucleotide encoding a human
CC      citron rho/rac-interacting kinase polypeptide. The isolated
CC      polynucleotide comprises a 6165 or 6603 base pair sequence, given in the
CC      specification. The human citron rho/rac-interacting kinase (CRK)
CC      polypeptide and polynucleotide are useful in preventing, ameliorating, or
CC      treating diseases associated with human CRK dysfunction such as obesity
CC      and obesity-associated comorbidities (e.g. hypertension, coronary artery
CC      disease, hyperlipidemia, stroke, gout, osteoarthritis, some types of
CC      cancer including endometrial, breast, prostate and colon cancer),
CC      anorexia, cachexia, bulimia, central nervous system disorders (e.g. mood
CC      disorders, anxiety disorders, Parkinson's disease or Alzheimer's
CC      disease), chronic obstructive pulmonary disease, or diabetes. These can
CC      also be used to treat pain associated with the disorders. The human CRK
CC      polypeptide is also useful in diagnostic assays or in genetic testing.
CC      The expression vector or the reagent is useful in preparing a medicament
CC      for modulating the activity of a human CRK in a disease, e.g. obesity, a
CC      central nervous system disorder, or chronic obstructive pulmonary
CC      disease. The fusion protein is useful for generating antibodies against a
CC      CRK polypeptide and for use in various assay systems. The methods are
CC      useful in producing and detecting the polynucleotide and polypeptide and
CC      in screening for agents that modulate the activity of the human CRK
CC      polypeptide. This sequence represents a protein relating to the human
CC      CRK protein of the invention
XX
SQ      Sequence 2055 AA;
XX
Query Match      96.3%; Score 10109.5; DB 6; Length 2055;
Best Local Similarity 96.0%; Pred. No. 0;
Matches 1974; Conservative 40; Mismatches 39; Indels 3; Gaps 2;
Qy      1 MLKFKYGARNPLDAGAAPIASRASRLNLFQCKPPTMQOQMSPLREGILDAFLVFE 60
Db      1 MLKFKYGVNRNPPASASBPISASRSRLNLFQCKPPTMQOQMSPLREGILDAFLVFE 60
Qy      61 ECSQPALMKIKHVSNFVKRYSDTTLAEQLQPSAKDEFVSLVCGHFAEYQVVRKATG 120
Db      61 ECSQPALMKIKHVSNFVKRYSDTTLAEQLQPSAKDEFVSLVCGHFAEYQVVRKATG 120
Qy      121 DIYAMKMKKKALLAQOVSPFEERNTILRSSTSPWIPQIYAFQDKNHLVWMEYQPGG 180
Db      121 DIYAMKMKKKALLAQOVSPFEERNTILRSSTSPWIPQIYAFQDKNHLVWMEYQPGG 180
Qy      181 DLSTLNRYEDQDENLQFYLAELIYAVSVHLMGVVHVDIKENILVDRTHIKLVDF 240
Db      181 DFLSTLNRYEDQDENLQFYLAELIYAVSVHLMGVVHVDIKENILVDRTHIKLVDF 240
Qy      241 GSAAKNSNMVNNAKPIGTDPYAPPEVLTVMNGDGKTYGLDQDMSVGIAYAYEM1YGR 300
Db      241 GSAAKNSNMVNNAKPIGTDPYAPPEVLTVMNGDGKTYGLDQDMSVGIAYAYEM1YGR 300
Qy      301 SPFAEGTSATFPNNINMFORFLKPPDDPKVSSDFLDLQSLCGQKRLKFBGICCHPFF 360
Db      301 SPFAEGTSATFPNNINMFORFLKPPDDPKVSSDFLDLQSLCGQKRLKFBGICCHPFF 360
Qy      361 SKIDMNNIRNSPPFVFTLSDDDTSNFDPEPKNSWSSPCQLSPGSGEELPFQFGS 420
Db      360 ARTDMMNIRNSPPFVFTLSDDDTSNFDPEPKNSWSSPCQLSPGSGEELPFQFGS 420
Qy      421 YSKALGILGRSESVSGLDSPATSSMEKLLIKSKELQDSQKCHMEQEMTRLHRVS 480
Db      420 YSKALGILGRSESVSGLDSPATSSMEKLLIKSKELQDSQKCHMEQEMTRLHRVS 480
Qy      481 EYEAVALSQKEVELKASERTQSLLEQDLATYITTECSSIKRLEQARMEVSGEDKALQLH 540
Db      480 EYEAVALSQKEVELKASERTQSLLEQDLATYITTECSSIKRLEQARMEVSGEDKALQLH 540

```

QY 541 DIBQSKLOEIKQEOVQAOVEKRLMMNOLBEDLVASRRRSDLYESSELSRLAAEFK 600
 DB 540 DIRQSKLOEIKQEOVQAOVEKRLMMNOLBEDLVASRRRSDLYESSELSRLAAEFK 599
 QY 601 RKATECCHKLKAKDOQPEVGEYAKLEKINAEQOLIOELOELKFAVASTATELLO 660
 DB 600 RKANECCHKLKAKDOQPEVGEYSKLEKINAEQOLIOELOELKFAVASTATELLO 659
 QY 661 NIRAQERARELEKLONRDESSGIRKLVAEERHSLNNKVKRLTEMRERENRLKDD 720
 DB 660 NIRAQERARELEKLONRDESSGIRKLVAEERHSLNNKVKRLTEMRERENRLKDD 719
 QY 721 IQTSQSOIQOMADKILBLEEKHRAQVSAQHLFVHLKQKHVEBEKI KVLDNQIKKDLAD 780
 DB 720 IQTSQSOIQOMADKILBLEEKHRAQVSAQHLFVHLKQKHVEBEKI KVLDNQIKKDLAD 779
 QY 781 KETLENNMORHEEBAHKGKILSEOKAMINAMDSKISLEORIVELSEANKLAANSLSFT 840
 DB 780 KESLENNMORHEEBAHKGKILSEOKAMINAMDSKISLEORIVELSEANKLAANSLSFT 839
 QY 841 QRNMKAQEMISELRQOKFYLETQAGKLEAQNRLKEQLEKISHQDHSKXNRLLEETRL 900
 DB 840 QRNMKAQEMISELRQOKFYLETQAGKLEAQNRLKEQLEKISHQDHSKXNRLLEETRL 899
 QY 901 REVSLHEBEOKLEKROLTELOLSQRESQUTLQAPALAEBSQLRQAKTELETTAAE 960
 DB 900 REVSLHEBEOKLEKROLTELOLSQRESQUTLQAPALAEBSQLRQAKTELETTAAE 959
 QY 961 EEEIOALTARHDEIQRFPDALRNSCTVITDLEBOLNLTEDNALNNQNFILSKQLEAS 1020
 DB 960 EEEIOALTARHDEIQRFPDALRNSCTVITDLEBOLNLTEDNALNNQNFILSKQLEAS 1019
 QY 1021 GANDEIVOLRSEVDHLREITEREMOLTSQKQTEALKTCTMLEBQVMDLEALNDELLE 1080
 DB 1020 GANDEIVOLRSEVDHLREITEREMOLTSQKQTEALKTCTMLEBQVMDLEALNDELLE 1079
 QY 1081 KERQWEMRBSVLGDEKQFECGRVRELORMIDTEKQSAARADQRTTESQVVELAVKXHA 1140
 DB 1080 KERQWEMRBSVLGDEKQFECGRVRELORMIDTEKQSAARADQRTTESQVVELAVKXHA 1139
 QY 1141 EITLALQOALKEOKKASLSDPKINDLEKKAEMENARSLQOKLETRELEKORLLEBOAK 1200
 DB 1140 EITLALQOALKEOKKASLSDPKINDLEKKAEMENARSLQOKLETRELEKORLLEBOAK 1199
 QY 1201 LQOQMDLQKNHIFELTQGLQEALDRADLKTERSDLEYOLENIOVLYSHEKVXMEGTISQ 1260
 DB 1200 LQOQMDLQKNHIFELTQGLQEALDRADLKTERSDLEYOLENIOVLYSHEKVXMEGTISQ 1259
 QY 1261 QTKLIDFLQAKMDQPAKKKVPLOYNELKALAEKAKCALEBALQKTRIELRSABEEA 1320
 DB 1260 QTKLIDFLQAKMDQPAKKKVPLOYNELKALAEKAKCALEBALQKTRIELRSABEEA 1319
 QY 1321 AHRKATDHPHSTPATAROQIAMSATVRSPEHOPSAMSLAPSSRRKESSTPEEFSRRL 1380
 DB 1320 AHRKATDHPHSTPATAROQIAMSATVRSPEHOPSAMSLAPSSRRKESSTPEEFSRRL 1379
 QY 1381 KERMHNNI PHRFNVGLMRAATKCAVCLDTVFHFGQASKCLEQCYMCHPKCSTCLPATCGL 1440
 DB 1380 KERMHNNI PHRFNVGLMRAATKCAVCLDTVFHFGQASKCLEQCYMCHPKCSTCLPATCGL 1439
 QY 1441 PAEYATHTFEAFCDKKNNSPGLQTKEPSSSLHLEGMKVPVPRNKRGGQGMDRKTYVLBGS 1500
 DB 1440 PAEYATHTFEAFCDKKNNSPGLQTKEPSSSLHLEGMKVPVPRNKRGGQGMDRKTYVLBGS 1499
 QY 1501 KVLITYNDEARBAGORPVEEPELCPDGDVSIHGAVGASELANTAKADVPTILKMHSHPH 1560
 DB 1500 KVLITYNDEARBAGORPVEEPELCPDGDVSIHGAVGASELANTAKADVPTILKMHSHPH 1559
 QY 1561 TCWPGRTLYLLAPSPDKORVVTALBESVVAAGRVSREKALEADAKILGNLSLKLJEGDRLD 1620
 DB 1560 TCWPGRTLYLLAPSPDKORVVTALBESVVAAGRVSREKALEADAKILGNLSLKLJEGDRLD 1619
 QY 1621 MNCITLPPSDOVVLVGTBEGLYALANVLKNSLTHPGIGAFQIYI IKOLEKULMAGEERA 1680

DB 1620 MNCITLPPSDOVVLVGTBEGLYALANVLKNSLTHPGIGAFQIYI IKOLEKULMAGEERA 1679
 QY 1681 LCLVDVKYKQSLAOSHLPAPQDISPNIPEAVKGCHEFGAKIENGCLICGAMPKRVYL 1740
 DB 1680 LCLVDVKYKQSLAOSHLPAPQDVSPNIFEA VVGCHLPAGKIENSLCICGAMPKRVYL 1739
 QY 1741 RYNNENSKTCIRKEITSEPCSCIHFTNYSILIGTKKFEYIDMKOYTLBFLDKNDHSLA 1800
 DB 1740 RYNNENSKTCIRKEITSEPCSCIHFTNYSILIGTKKFEYIDMKOYTLBFLDKNDHSLA 1799
 QY 1801 PAVFAASNSFPVSIQVNSAGOREEYLLCFHEFGVFDVSGRRSTDDLKMGRLPLAFA 1860
 DB 1800 PAVFAASNSFPVSIQVNSAGOREEYLLCFHEFGVFDVSGRRSTDDLKMGRLPLAFA 1859
 QY 1861 YREPYLFVTHFNLSLEYIEIQARSSACTPARAYLDIPNPRYLGAISSGAIYLAASYODKL 1920
 DB 1860 YREPYLFVTHFNLSLEYIEIQARSSACTPARAYLEIPNPRYLGAISSGAIYLAASYODKL 1919
 QY 1921 RYVCCGKLVKESGTEHHGSPSTSRSSPNKRGPTTNEHTTKRVASSPAPBEGSPHREP 1980
 DB 1920 RYVCCGKLVKESGTEHHGSPSTSRSSPNKRGPTTNEHTTKRVASSPAPBEGSPHREP 1979
 QY 1981 STEPHRY--RREGRTLRDSDSPGRPLERKSPGRILSTRERSPARLFEDESRGRPLAGAV 2038
 DB 1980 STEPHRYRDRGRTELRDSDSPGRPLERKSPGRILSTRERSPARLFEDESRGRPLAGAV 2039
 QY 2039 RTPLSQNVKVMQSSV 2054
 DB 2040 RTPLSQNVKVMQSSV 2055

RESULT 15
 AAE24079
 ID AAE24079 standard; protein; 2053 AA.
 XX
 AC AAE24079;
 XX
 DT 04-OCT-2002 (first entry)
 XX
 DE Human MDPK protein.
 XX
 KW Human; myotonic dystrophy type protein kinase; MDPK; 13245 protein;
 KW tumorigenesis; tumour growth; tumour metastasis; viral infection;
 KW skeletal muscle disorder; muscular dystrophy; myotonic dystrophy;
 KW immune disorder; neoplastic disorder; gene therapy.
 OS Homo sapiens.
 XX
 FH Key
 FT 53..303 Location/Qualifiers
 FT Domain /note="pkinase domain"
 FT 97..360 /note="pkinase domain"
 FT 195..210 /note="pkinase domain"
 FT Region /note="Antigenic epitope"
 FT 217..229 /note="Serine/Threonine protein kinase active site
 FT signature"
 FT 455..475
 FT Region /note="Antigenic epitope"
 FT 1568..1865
 FT Domain /note="CNH domain"
 PN WO200234896-A2.
 XX
 PD 02-MAY-2002.
 XX
 PF 23-OCT-2001; 2001WO-US050636.
 XX
 PR 23-OCT-2000; 2000US-0242429P.
 XX
 PA (MILL-) MILLENNIUM PHARM INC.

XX Kapeller-Libermann R;
 XX WPI; 2002-479720/51.
 DR N-PSDB; AAD39191.
 XX
 XX Human myotonic dystrophy type protein kinase polypeptide and
 PT polynucleotide useful for prognosticating, diagnosing, preventing or
 PT inhibiting tumorigenesis, tumor growth, tumor metastasis and viral
 PT infection.
 XX
 XX Claim 8; Fig 3; 148pp; English.
 XX
 XX The invention relates to human myotonic dystrophy type protein kinase
 CC (MDPK) polypeptides designated as 13245 and nucleic acid molecules
 CC encoding such polypeptides. 13245 molecules are used to develop
 CC diagnostic and therapeutic agents for prognosticating, diagnosing,
 CC preventing, inhibiting, alleviating or curing MDPK-related disorders.
 CC Polypeptides of the invention are used to develop diagnostic and
 CC therapeutic agents for 13245-mediated or related disorders such as
 CC tumorigenesis, tumor growth, tumor metastasis, viral infection of a
 CC cell, skeletal muscle disorders (e.g. muscular and myotonic dystrophies),
 CC immune disorders and neoplastic disorders. The invention is also used in
 CC gene therapy. The present sequence is human MDPK protein
 XX
 SQ Sequence 2053 AA;
 Query Match 95.5%; Score 10022.5; DB 5; Length 2053;
 Best Local Similarity 95.7%; Pred. No. 0;
 Matches 1979; Conservative 3; Mismatches 5; Indels 81; Gaps 3;

DB 601 RKATECQHKLKADQKPEVGEYAKLEKINAEOQLKIOELQELKEXA----- 648
 QY 661 NIROAKERAERELEKQNRDESSGQIRKCLVEAEERRSLNNKYRLTETMERENRLKXD 720
 DB 649 ----AKERAERELEKQNRDESSGQIRKCLVEAEERRSLNNKYRLTETMERENRLKXD 704
 QY 721 IQTSQOIQQWADKILILEEKGREAOVSAGHLEHNLKKEOHYEKIVLNDQIKDLAD 780
 DB 705 IQTSQOIQQWADKILILEEKGREAOVSAGHLEHNLKKEOHYEKIVLNDQIKDLAD 764
 QY 781 KETLENNMQREEREAHEKGLTSEOKAMINMDKIRSLERIVELSPANKLAANSLSFT 840
 DB 765 KETLENNMQREEREAHEKGLTSEOKAMINMDKIRSLERIVELSPANKLAANSLSFT 824
 QY 841 QRNKAQEMISELRQOKFYLETQAGLEAQNRKLEOLEKISHQSDSKNRLLELETRL 900
 DB 825 QRNKAQEMISELRQOKFYLETQAGLEAQNRKLEOLEKISHQSDSKNRLLELETRL 884
 QY 901 REVSLHEBQKLEIKROULTILOSLQERESQITLQAARALLESQLRQAKTELETTAAE 960
 DB 885 REVSLHEBQKLEIKROULTILOSLQERESQITLQAARALLESQLRQAKTELETTAAE 944
 QY 961 EEEIOALTARHDEIQRKFDALRNSCTVITDLEEQNLQTEONABLNNQNFYLSQLOEAS 1020
 DB 945 EEEIOALTARHDEIQRKFDALRNSCTVITDLEEQNLQTEONABLNNQNFYLSQLOEAS 1004
 QY 1021 GANDEIVOLNSEVVDLREITEREMQITSQKOTMEALKTCTMEBOQMDLEALNDELLE 1080
 DB 1005 GANDEIVOLNSEVVDLREITEREMQITSQKOTMEALKTCTMEBOQMDLEALNDELLE 1064
 QY 1081 KERQWEAMRSVLDGKESQFECEVRELQPMIDTEKOSPARADQRTTESQVVELAVKEKA 1140
 DB 1065 KERQWEAMRSVLDGKESQFECEVRELQPMIDTEKOSPARADQRTTESQVVELAVKEKA 1124
 QY 1141 EILALQALKEQKKAASLSKNDLEKKAMLENNARSIOQKLETRERIKORLLEBAK 1200
 DB 1125 EILALQALKEQKKAASLSKNDLEKKAMLENNARSIOQKLETRERIKORLLEBAK 1184
 QY 1201 LQQQMDLQKNIIFRLTQLOEALDRADLKTERRSDLEYOLENIOVLSHEKYKKEGTSQ 1260
 DB 1185 LQQQMDLQKNIIFRLTQLOEALDRADLKTERRSDLEYOLENIOVLSHEKYKKEGTSQ 1244
 QY 1261 QTKLIDFLQAKMDQPAKXXX-----VPLQYNELKALKEKARCAELEEA 1305
 DB 1245 QTKLIDFLQAKMDQPAKXXX-----VPLQYNELKALKEKARCAELEEA 1304
 QY 1306 LQKTRILRSAREEAAARKATDHPHSTPATARQOIMSAIVSPEHQPMSMLAPSS 1365
 DB 1305 LQKTRILRSAREEAAARKATDHPHSTPATARQOIMSAIVSPEHQPMSMLAPSS 1364
 QY 1366 RRKESSTPEEFSSRLKERMHNNIPIHFNVGILMRAATCAVCLDTVHGRQASKLECOVM 1425
 DB 1365 RRKESSTPEEFSSRLKERMHNNIPIHFNVGILMRAATCAVCLDTVHGRQASKLECOVM 1424
 QY 1426 CHPKCSTCLPATGCLPAEYATHTFEAFCRDKANSPGLOTEPSSSLHLEGMMKVPNNKR 1485
 DB 1425 CHPKCSTCLPATGCLPAEYATHTFEAFCRDKANSPGLOTEPSSSLHLEGMMKVPNNKR 1484
 QY 1486 GQGWDRKXIVLBSKYLIDNEAREAGORPVEFEELCPDGDVSIGGAVGASGLANTAK 1545
 DB 1485 GQGWDRKXIVLBSKYLIDNEAREAGORPVEFEELCPDGDVSIGGAVGASGLANTAK 1544
 QY 1546 ADVPYILKJMSHPHTTGWPGRTLYLAPSPDDQRWVTALBSVAGRVREKXEAADAKL 1605
 DB 1545 A-----EKXEAADAKL 1554
 QY 1606 LQNSLLKLEGGDRLDNMNCTLPFSDQVVLVGTBEGLYALANVLSKSLTHVPGIGAVFOIYII 1665
 DB 1555 LQNSLLKLEGGDRLDNMNCTLPFSDQVVLVGTBEGLYALANVLSKSLTHVPGIGAVFOIYII 1614
 QY 1666 KDLKELMTLIGERBALCLVDYKVKVOSLAOSHLPAQPDISPNIPEAVKGCHLFGAGKIN 1725
 DB 1615 KDLKELMTLIGERBALCLVDYKVKVOSLAOSHLPAQPDISPNIPEAVKGCHLFGAGKIN 1674

```
QY 1726 GLCICAMPSKVILRYNENLSKYCIKRIETSEPCSIHFTNYSILIGTNKFEYIDMKQ 1785
Db 1675 GLCICAMPSKVILRYNENLSKYCIKRIETSEPCSIHFTNYSILIGTNKFEYIDMKQ 1734
QY 1786 YTLIEEFLDKNDHSLAPAVFAASSNSFPVSIQVNSAQOREEYLLCFHEGCVFVDSYGRRS 1845
Db 1735 YTLIEEFLDKNDHSLAPAVFAASSNSFPVSIQVNSAQOREEYLLCFHEGCVFVDSYGRRS 1794
QY 1846 RTDDLKMSRLPLAFAYREPYLFTTHFNSLEVIHQARSSAGTPARAYLDIPNPRYLGPAL 1905
Db 1795 RTDDLKMSRLPLAFAYREPYLFTTHFNSLEVIHQARSSAGTPARAYLDIPNPRYLGPAL 1854
QY 1906 SSGAIYLAASYODKLRYICCKGNLVKESGTEHHRGPSSTRSSPNKRGPPTYNEHITKRA 1965
Db 1855 SSGAIYLAASYODKLRYICCKGNLVKESGTEHHRGPSSTRSSPNKRGPPTYNEHITKRA 1914
QY 1966 SSPAPPEGPGSHPREPSTPHRYREGRTELRDKSPGRPLEREKSPGRILSTRERSPARLF 2025
Db 1915 SSPAPPEGPGSHPREPSTPHRYREGRTELRDKSPGRPLEREKSPGRMLSTRERSPGRLF 1974
QY 2026 EDSRRGRLPAGAVRTPLSQVNKXWDOSS 2053
Db 1975 EDSRRGRLPAGAVRTPLSQVNKXWDOSS 2002
```

Search completed: May 11, 2005, 16:40:19
Job time : 239 secs